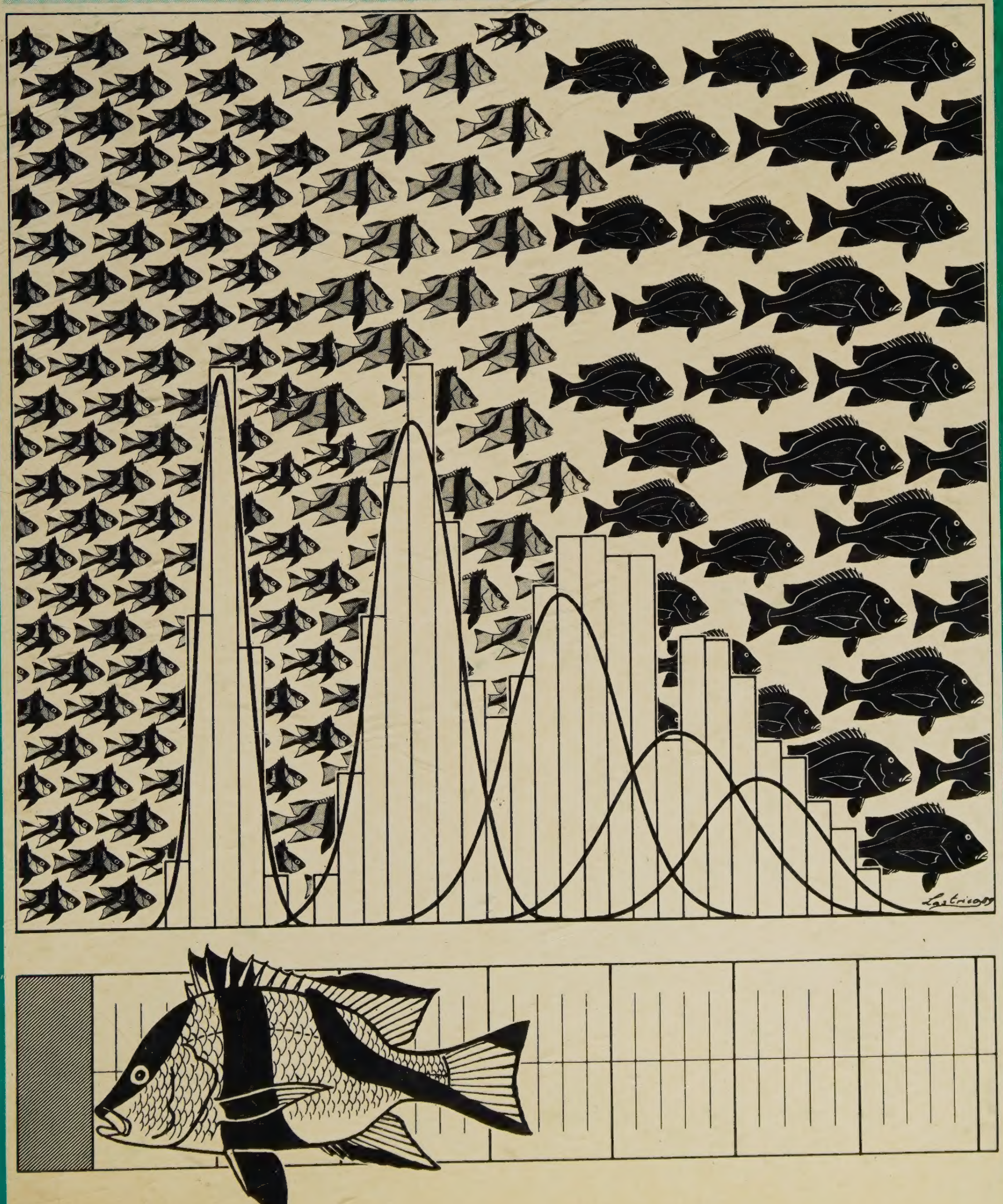
Introduction to tropical fish stock assessment Part 1 - Manual

FAO
FISHERIES
TECHNICAL
PAPER

306/1



DANIDA

DENMARK
FUNDS-IN-TRUST
GCP/INT/392/DEN

FOOD
AND
AGRICULTURE
ORGANIZATION
OF THE
UNITED NATIONS

Digitized by the Internet Archive
in 2025

https://archive.org/details/fao-fisheries-and-aquaculture-technical-paper_1989_306-1

Introduction to tropical fish stock assessment
Part 1 - Manual

by
Per Sparre
Marine Resources Service
Fishery Resources and Environment Division
FAO Fisheries Department
Rome, Italy

Erik Ursin
Consultant
GCP/INT/392/DEN
Gislinge, Denmark

Siebren C. Venema
Project Manager
GCP/INT/392/DEN
Rome, Italy

FAO FISHERIES TECHNICAL PAPER

306/1

DANIDA

**DENMARK
FUNDS-IN-TRUST
GCP/INT/392/DEN**

**FOOD
AND
AGRICULTURE
ORGANIZATION
OF THE
UNITED NATIONS
Rome, 1989**

The designations employed and the presentation of material in this publication do not imply the expression of any opinion whatsoever on the part of the Food and Agriculture Organization of the United Nations concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

M-43
ISBN 92-5-102850-8

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise, without the prior permission of the copyright owner. Applications for such permission, with a statement of the purpose and extent of the reproduction, should be addressed to the Director, Publications Division, Food and Agriculture Organization of the United Nations, Via delle Terme di Caracalla, 00100 Rome, Italy.

© **FAO 1989**

PREPARATION OF THIS DOCUMENT

The first edition of this manual was prepared by the FAO/DANIDA project "Training in fish stock assessment and fisheries research planning" (GCP/INT/392/DEN) for use in a series of regional and national training courses on fish stock assessment organized by the project from 1983 to early 1989.

In 1984 the author Per Sparre, then employed by the Danish Institute of Fisheries and Marine Research, Charlottenlund, Denmark, was asked to write this manual on the basis of lecture notes and case studies prepared by the following team of lecturers engaged in the courses:

Mr. Peder Agger of Roskilde University, Denmark, Dr. Ole Bagge, Dr. Jan Beyer, Mr. Jørgen Møller Christensen, Mr. Poul Degnbol, Mr. Henrik Gislason, Mr. Eskild Kirkegaard, Mr. Hans Lassen, Mr. Niels A. Nielsen, Mr. Per Sparre and Dr. Erik Ursin of the Danish Institute for Fisheries and Marine Research, Dr. Daniel Pauly of the International Center for Living Aquatic Resources Management and Mr. Siebren C. Venema of FAO.

The first edition was printed in July 1985 in Manila, Philippines, and distributed by the project through the Network of Tropical Fisheries Scientists of ICLARM and training courses.

The demand for copies of the manual turned out to be high and in 1986 an FAO/DANIDA project evaluation team strongly recommended it to be edited, reprinted and distributed widely.

In the mean time the project has organized three follow-up courses, where participants were assisted in preparing a scientific paper based on their own data sets, respectively in Hirtshals, Denmark (1986), Manila, the Philippines (1987) and Cochin, India (1987). These courses exposed the need for an expansion of the manual to cover assessments of migratory stocks and gear selectivity and the modification of other subjects.

Further a large number of useful comments and suggestions were received from various users of the first edition in particular the participants and local staff of the courses organized between 1985 and 1989, Dr. E.L. Olsen of the University of Papua New Guinea, Dr. S. Garcia, Chief of the Marine Resources Service, FAO and many others.

Dr. Keith Brander (Fisheries Laboratory, Lowestoft, UK) and Mr. Niek van Zalinge (FAO, Rome) were engaged by the project to review the manual, including some of the additions and they also came up with many useful suggestions and corrections.

Early 1989 it was decided to make a more thorough review of the manual and to publish it as a FAO Fisheries Technical Paper. Dr. Erik Ursin (consultant) and Siebren C. Venema (project manager) of the FAO/DANIDA project, undertook this revision in close cooperation with Per Sparre (now Fishery Resources Officer, Marine Resources Service, FAO, Rome). The figures, of which the originals were all drawn by Per Sparre, were partly revised and redrawn by Mr. Oliviero Lidonnici and a new cover was designed by Per Sparre and Paolo Lastrico (FAO). The references were reviewed by Ms. Gloria Soave and Siebren Venema and last, but not least, the text was completely retyped and rearranged by Ms. Jane Ugilt (Denmark).

Due to time constraints the text has not been reviewed by an English mother tongue editor, and we hope that you will forgive us for the occasional odd wording. Translations into French (already partly completed on the basis of the first edition by the International Center for Ocean Development, Canada), Spanish and Portuguese, are scheduled to appear in the course of 1990 or early 1991.

Sparre, P.; Ursin, E.; Venema, S.C.
Introduction to tropical fish stock assessment.
Part 1. Manual.
FAO Fisheries Technical Paper. No. 306.1. Rome, FAO.
1989. 337p.

ABSTRACT

A selection of methods on fish stock assessment is described in detail, with examples of calculations. Special emphasis is placed on methods applicable to tropical fisheries, i.e. those based on the analysis of length-frequencies. After a short introduction to statistics, it covers the estimation of growth parameters and mortality rates, virtual population methods (including length and age based methods), gear selectivity, the yield per recruit model of Beverton and Holt, the estimation of maximum sustainable yield using surplus production models, multispecies and multifleet problems, the assessment of migratory stocks, a discussion on the stock/recruitment relationship and the execution of demersal trawl surveys. The link between this manual and the Length Frequency Stock Assessment (LFSA) package of microcomputer programs is explained. The manual is completed with an extensive list of references (including material for further reading), a list of symbols used, a subject index and a number of exercises with solutions.

Distribution:

DANIDA
Participants at FAO/DANIDA courses
on Fish Stock Assessment
Members of the Network of Tropical
Fisheries Scientists
Institutes specialised in Tropical
Fish Stock Assessment
Institutes of Fisheries Education
Interested National and International Organizations
Marine and Inland Selectors
FAO Regional Fishery Officers
FAO Fisheries Department
GCP/INT/392/DEN

CONTENTS

Page

PART 1 MANUAL

1 INTRODUCTION 1

1.1 The primary objective of fish stock assessment 1

1.2 The stock concept 2

1.3 Models 5

1.3.1 Analytical models 7
1.3.2 Holistic models 9

1.4 Assessment of stocks in tropical waters 10

1.5 Definitions of body length 12

1.6 Age and recruitment 13

1.7 The underlying assumption of random samples 14

1.8 The organization of the manual 15

1.9 Further reading 19

2 BIOSTATISTICS 20

2.1 Mean value and variance 20

2.2 The normal distribution 23

2.3 Confidence limits 28

2.4 Ordinary linear regression analysis 32

2.5 The correlation coefficient and functional regression 37

2.6 Linearizing transformations 41

2.7 Simple random sampling 47

2.8 Stratified random sampling 51

3 ESTIMATION OF GROWTH PARAMETERS 57

3.1 The von Bertalanffy growth equation 57

3.1.1 Variability and applicability of growth parameters 61

3.2 Input data for the von Bertalanffy growth equation 62

3.2.1 Data from age readings and length measurements 63
3.2.2 Length composition data (without age compositions) 67

3.2.2.1 Commercial catches 68

3.3 Estimation of growth parameters from length at age data 68

3.3.1 The Gulland and Holt plot 68
3.3.2 The Ford-Walford plot 71

Page

3.3.3 The von Bertalanffy plot 73
3.3.4 The least squares method 75
3.3.5 The use of age/length keys 75
3.3.6 The weight converted von Bertalanffy growth equation 77

3.4 Estimation of age composition from length frequencies 79

3.4.1 Bhattacharya's method 84

3.4.1.1 Basis of the computation procedure of the Bhattacharya method 84
3.4.1.2 The computation procedure of the Bhattacharya method 85
3.4.1.3 More features of the Bhattacharya analysis 97

3.4.2 Modal progression analysis 98
3.4.3 The probability paper method and the parabola method 107

3.5 Computer based length frequency analysis 109

3.5.1 ELEFAN I 109
3.5.2 The seasonalized von Bertalanffy growth equation 115
3.5.3 Maximum likelihood methods 115
3.5.4 Limitations of length frequency analysis 119

3.6 The use of data on gonadal maturity 122

3.7 Comparison of growth curves, Munro's phi prime 122

4 ESTIMATION OF MORTALITY RATES 124

4.1 The concept of a cohort and some basic notation 124

4.2 The dynamics of a cohort 126

4.3 Estimating Z from CPUE data and the concept of the catchability coefficient 132

4.3.1 Heincke's method 136

4.4 Estimation of Z from a linearized catch curve 137

4.4.1 The constant parameter system 137
4.4.2 The linearized catch equation 139
4.4.3 The linearized catch curve based on age composition data 141
4.4.4 The linearized catch curve based on age compositions with variable time intervals 142
4.4.5 The linearized catch curve based on length compostion data 144
4.4.6 The cumulated catch curve based on length composition data. (The Jones and van Zalinge method) 148
4.4.7 Summary of the linearized catch curve methods 150

4.5 Beverton and Holt's Z-equations 150

4.5.1 Beverton and Holt's Z-equation based on length data 153
4.5.2 Beverton and Holt's Z-equation based on age data 154

Page

4.5.3 Beverton and Holt's Z-equation based on length at first capture 155
4.5.4 The Wetherall et al. method 156

4.6 A plot of Z on effort for separate estimates of F and M 158

4.7 Natural mortality 160

4.7.1 Natural mortality and longevity 161
4.7.2 Pauly's empirical formula 162
4.7.3 Rikhter and Efanov's formula 163

5 VIRTUAL POPULATION METHODS 164

5.1 Virtual population analysis (VPA) 164

5.2 Age based cohort analysis (Pope's cohort analysis) 169

5.3 A predictive age based model (Thompson and Bell) 173

5.4 Jones' length based cohort analysis 179

5.5 A predictive length based model (Thompson and Bell) 185

6 GEAR SELECTIVITY 192

6.1 Trawl selection 192

6.1.1 A mathematical model for the selection ogive 195
6.1.2 Estimation of the selection ogive from a catch curve 198
6.1.3 Mesh selection experiments 202

6.2 Estimation of gill net selection 204

6.3 Discussion of selection by other gears 213

6.4 Estimation of selection curves from cohort analysis 215

6.5 Using a selection curve to adjust catch samples 215

7 SAMPLING COMMERCIAL CATCHES 219

7.1 An example of data collection 219

7.2 Estimation of total catch in weight 221

7.3 Estimation of total length composition 223

8 THE YIELD PER RECRUIT MODEL OF BEVERTON AND HOLT 231

8.1 The exponential decay model behind the Beverton and Holt model 233

8.2 Yield per recruit 234

8.3 Biomass per recruit 239

8.4 Beverton and Holt's relative Y/R concept 241

Page

8.5 Yield per recruit from length data 242

9 ESTIMATION OF MAXIMUM SUSTAINABLE YIELD USING SURPLUS PRODUCTION MODELS 243

9.1 The Schaefer model and the Fox model 243

9.2 The assumptions behind the surplus production models 248

9.3 Gulland's empirical formula 249

9.4 Cadima's formula 250

9.5 MSY estimates based on the surplus production model 252

9.6 Validation of MSY-estimates based on empirical formulas 253

9.7 Munro and Thompsons's plot 254

9.8 Standardization of effort 255

9.9 The Deriso/Schnute delay difference model 259

10 MULTISPECIES/MULTIFLEET PROBLEMS 261

10.1 Surplus production models applied to a multispecies/multifleet system 261

10.2 Biological interaction 263

10.3 Economic interaction 264

10.4 Technical interaction 264

10.4.1 A yield per recruit model for mixed fisheries 264
10.4.2 Assessment of mixed fisheries based on length frequency data 265
10.4.3 Multifleet mixed fisheries 267

11 ASSESSMENT OF MIGRATORY STOCKS 271

11.1 The concept and study of migration 271

11.2 Bias caused by migration 274

11.3 The annual-return matched samples method 280

11.3.1 Growth parameters estimated by the annual-return matched samples method 280

11.4 The general matched samples method 284

11.5 Assessment based on tagging data 286

11.6 Estimation of the growth parameters of a migratory stock: the Atlantic mackerel 287

12 THE STOCK/RECRUITMENT RELATIONSHIP 290

12.1 Classical S/R considerations 293

Page

12.2 The stability of recruitment 294
12.3 Towards modeling recruitment 295

13 DEMERSAL TRAWL SURVEYS 297

13.1 The bottom trawl 297

13.2 Planning a demersal trawl survey 298

13.3 Data recording 299

13.4 Deck sampling and catch recording procedures 300

13.5 The swept area 302

13.6 Biomass estimation by the swept area method 303

13.7 Precision of the estimate of biomass 304

13.8 Estimation of maximum sustainable yield 306

14 LFSA PACKAGE OF MICROCOMPUTER PROGRAMS 307

14.1 Length frequency (LF) programs 307

14.2 Age/length (AL) analysis: estimation of growth parameters from age/length data 310

14.3 Various other programs 311

15 REFERENCES 312

16 SUBJECT INDEX 333

PART 2

17 EXERCISES

18 SOLUTIONS TO EXERCISES

LIST OF SYMBOLS USED

A: Mathematical notation (general)

$*$	multiplication sign
/	division
ln	natural logarithm (base e = 2.7182818)
log	10 based logarithm
exp	exponential function, $\exp(x) = e^{x}$
$\sum_{i=1}^{i=n} X(i)$	sum of all values of X(i), for i from 1 to n; the sum X(1)+X(2)+...+X(n)
$\surd$ or $\sqrt{}$	square root
∞	infinity
Δx	delta x, a small increment of the variable x
$\underset{j}{\mathrm{MAX}} \{ X(j) \}$	maximum value among the elements in the set { X(j) } = { X(1),X(2),...X(j),...)
$\bar{x}$	mean value of x
x(i,j)	i,j indices of x (usually printed as $x_{i,j}$)
π	pi = 3.14159
a < b	a smaller than b
a > b	a greater than b
a => b	a greater than or equal to b

B. Statistical notation

		First introduced Section
linear regression y = a + bx		2.4
a	intercept	-
b	slope	-
x	independent variable	-
y	dependent variable	-
r	correlation coefficient	-
a'	intercept of functional regression	2.5
b'	slope of functional regression	-
n	number of observations	2.1
$\bar{x}$	mean value of x	2.1
s^2	variance	
s	standard deviation	2.1
$s/\bar{x}$	relative standard deviation or coefficient of variation	2.1
$s/\sqrt{n}$	standard error	2.1
t(f)	fractile of Student's t-distribution	2.3
f	degrees of freedom	2.3
ϵ	(epsilon) maximum relative error	2.7

C. Symbols used in formulas for fish stock assessment

		Section
a	swept area (effective path swept by a trawl)	13.5
ASP	available sum of peaks (ELEFAN)	3.5
b	constant in length-weight relationship $W = qL^b$	2.6
B	biomass	5.3
Bv	virgin (unexploited) biomass	8.3, 9.1
B/R	biomass per recruit	8.2
C	catch in numbers (VPA)	5.0
C(t,∞)	cumulated catch	4.4
C	amplitude (0-1) (ELEFAN)	3.5
CPUE	catch per unit of effort	4.3
D	number of natural deaths (VPA)	5.0
D50%	deselection, length at which 50% is not caught	6.2
dL	interval size of length	3.4
E	fishing effort	7.1
E	exploitation rate (F/Z)	8.4
ESP	explained sum of peaks (ELEFAN)	3.5
f	fishing effort (also as F)	9.0
F	fishing mortality coefficient or instantaneous rate (per time unit)	4.2
Fm	maximum fishing mortality	5.3
F-array	array of F at age, fishing pattern	5.1
I	separation index	3.5
K	curvature parameter	3.1
L	length	general
L1-L2	length class	-
L_∞ or L∞	L infinity, asymptotic length (mean length of very old fish)	3.1
L'	some length for which all fish of that length and larger are under full exploitation (lower limit of corresponding length interval)	4.5
Lc or L50%	length at which 50% of the fish is retained by the gear and 50% escape	4.5
L75% or L75	length at which 75% of the fish is retained in the gear	6.1
Lm	optimum length for being caught	6.2
m	= K/Z	8.4
M	natural mortality coefficient or instantaneous rate of natural mortality or natural mortality rate (per time unit)	4.2
MSE	Maximum Sustainable Economic Yield	5.5
MSY	Maximum Sustainable Yield	5.5
N	number of survivors (VPA)	5.0
N(t)	number of survivors of a cohort attaining age t	4.1
N(Tr)	number of recruits to the fishery	4.1
ϕ'	(phi prime), $\ln K + 2 * \ln L_\infty$	3.7
q	condition factor, constant in length-weight relationship	2.6
q	catchability coefficient	4.3
R	recruitment N(Tr)	4.1
S	selection curve	6.1
S(L)	logistic curve	6.1
S.F.	selection factor	6.1
S/R	stock recruitment relationship	12.0
t	time (usually in years)	general
T	ambient temperature in °C	4.7
Tc	age at first capture (start of exploited phase)	4.1
Tm	longevity (maximum age)	4.7
Tm 50	age of massive maturation (50% of population mature)	4.7
t_o	t-zero, initial condition parameter (in years)	3.1
Tr	recruitment age	4.1
ts	summerpoint (0-1) (ELEFAN)	3.5

LIST OF SYMBOLS USED

A: Mathematical notation (general)

*	multiplication sign
/	division
ln	natural logarithm (base e = 2.7182818)
log	10 based logarithm
exp	exponential function, $\exp(x) = e^x$
$\sum_{i=1}^{i=n} X(i)$	sum of all values of X(i), for i from 1 to n; the sum X(1)+X(2)+...+X(n)
√ or $\sqrt{}$	square root
∞	infinity
Δx	delta x, a small increment of the variable x
$\max_{j} \{ X(j) \}$	maximum value among the elements in the set { X(j) } = { X(1),X(2),...X(j),...)
$\bar{x}$	mean value of x
x(i,j)	i,j indices of x (usually printed as $x_{i,j}$)
π	pi = 3.14159
$a < b$	a smaller than b
$a > b$	a greater than b
a => b	a greater than or equal to b

B. Statistical notation

		First introduced Section
linear regression $y = a + bx$		2.4
a	intercept	-
b	slope	-
x	independent variable	-
y	dependent variable	-
r	correlation coefficient	-
a'	intercept of functional regression	2.5
b'	slope of functional regression	-
n	number of observations	2.1
$\bar{x}$	mean value of x	2.1
s^2	variance	
s	standard deviation	2.1
$s/\bar{x}$	relative standard deviation or coefficient of variation	2.1
$s/\sqrt{n}$	standard error	2.1
t(f)	fractile of Student's t-distribution	2.3
f	degrees of freedom	2.3
ϵ	(epsilon) maximum relative error	2.7

C. Symbols used in formulas for fish stock assessment

		Section
a	swept area (effective path swept by a trawl)	13.5
ASP	available sum of peaks (ELEFAN)	3.5
b	constant in length-weight relationship $W = qL^b$	2.6
B	biomass	5.3
Bv	virgin (unexploited) biomass	8.3, 9.1
B/R	biomass per recruit	8.2
C	catch in numbers (VPA)	5.0
$C(t,\infty)$	cumulated catch	4.4
C	amplitude (0-1) (ELEFAN)	3.5
CPUE	catch per unit of effort	4.3
D	number of natural deaths (VPA)	5.0
D50%	deselection, length at which 50% is not caught	6.2
dL	interval size of length	3.4
E	fishing effort	7.1
E	exploitation rate (F/Z)	8.4
ESP	explained sum of peaks (ELEFAN)	3.5
f	fishing effort (also as F)	9.0
F	fishing mortality coefficient or instantaneous rate (per time unit)	4.2
Fm	maximum fishing mortality	5.3
F-array	array of F at age, fishing pattern	5.1
I	separation index	3.5
K	curvature parameter	3.1
L	length	general
L1-L2	length class	-
L_∞ or L∞	L infinity, asymptotic length (mean length of very old fish)	3.1
L'	some length for which all fish of that length and larger are under full exploitation (lower limit of corresponding length interval)	4.5
Lc or L50%	length at which 50% of the fish is retained by the gear and 50% escape	4.5
L75% or L75	length at which 75% of the fish is retained in the gear	6.1
Lm	optimum length for being caught	6.2
m	= K/Z	8.4
M	natural mortality coefficient or instantaneous rate of natural mortality or natural mortality rate (per time unit)	4.2
MSE	Maximum Sustainable Economic Yield	5.5
MSY	Maximum Sustainable Yield	5.5
N	number of survivors (VPA)	5.0
N(t)	number of survivors of a cohort attaining age t	4.1
N(Tr)	number of recruits to the fishery	4.1
ϕ'	(phi prime), $\ln K + 2 * \ln L_\infty$	3.7
q	condition factor, constant in length-weight relationship	2.6
q	catchability coefficient	4.3
R	recruitment N(Tr)	4.1
S	selection curve	6.1
S(L)	logistic curve	6.1
S.F.	selection factor	6.1
S/R	stock recruitment relationship	12.0
t	time (usually in years)	general
T	ambient temperature in °C	4.7
Tc	age at first capture (start of exploited phase)	4.1
Tm	longevity (maximum age)	4.7
Tm 50	age of massive maturation (50% of population mature)	4.7
t_0	t-zero, initial condition parameter (in years)	3.1
Tr	recruitment age	4.1
ts	summerpoint (0-1) (ELEFAN)	3.5

tw	winterpoint (0-1) (ELEFAN)	3.5
t50%	age at which 50% of the fish is retained in the gear (Thompson and Bell)	(5.3) 6.1
U	1-Lc/L	8.4
v	value (Thompson and Bell)	5.3
VPA	Virtual Population Analysis	5.0
W∞	weight infinity, asymptotic weight (W infinity, mean weight of very old fish)	3.3
X factor	multiplication factor of F (Thompson and Bell)	5.3
y	yield	5.3
Y/R	yield per recruit	8.2
(Y/R)'	relative yield per recruit	8.4
Z	total mortality coefficient instantaneous rate of total mortality, total mortality rate (per time unit)	4.2

1 INTRODUCTION

There are several excellent textbooks and manuals dealing with fish stock assessment wherein the theory behind the various models and methods is well explained, including the mathematical derivation of formulas (see for example Gulland, 1969 and 1983 and Csirke, 1980a). The problem is that from such manuals the beginning fishery scientists can usually not derive the precise instructions needed to perform each analysis. Where experienced scientists are around such instructions can easily be provided on-the-job and through participation in working group meetings. However, there are still many countries where such a transfer is not possible. This manual is in particular intended to assist the beginning fishery scientists. It is an attempt to put down in writing that part of the instructions that is usually transferred through on-the-job training. It concentrates on the application of methods, while less attention is given to detailed explanations of the theory behind them.

With the help of this manual the fishery scientist should be able to make a start with data analyses and building up the necessary skills and insight in solving stock assessment problems. After this initial step it should be easier to access more complex textbooks.

Stock assessment of tropical resources has developed rapidly in the last decade in particular through the work of Pauly (1979, 1980), Saila and Roedel (1980), Pauly and David (1981), Garcia and Le Reste (1981) and Munro (1983), but also because of the rapid development of microcomputer hard- and software. This manual is intended to further contribute to this development and for that reason emphasis has been placed on those methods which are particularly useful in tropical areas, while most of the examples given are based on tropical stocks.

The rapid introduction of special software for fish stock assessment, in particular that based on length frequency data, such as the packages COMPLEAT ELEFAN (Gayanilo, Soriano and Pauly, 1988) and LFSA (Sparre, 1987) also means that inexperienced fishery scientists may be placed in a position where they are using methods and models without fully realising the limitations of each method. The present manual should provide the necessary background knowledge on the methods to the users of the software mentioned above, in particular the LFSA package. This does not mean that this manual is directly related to computers. On the contrary, every method and exercise can be applied with the help of a good, programmable scientific pocket calculator. Further details on the use of this manual in training courses can be found in Venema, Christensen and Pauly (1988a).

1.1 The primary objective of fish stock assessment

The basic purpose of fish stock assessment is to provide advice on the optimum exploitation of aquatic living resources such as fish and shrimp. Living resources are limited but renewable, and fish stock assessment may be described as the search for the exploitation level which in the long run gives the maximum yield in weight from the fishery.

Fig. 1.1.1 illustrates this basic objective of fish stock assessment. On the horizontal axis is the fishing effort measured, for example, in number of boat days fished. On the other axis is the yield, i.e. the landings in weight. (If the landings consist of different groups of animals, for example shrimp, finfish and squid, it may be more appropriate to express the yield in terms of value). It shows that up to a certain level we gain by increasing the fishing effort, but after that level the renewal of the resource (the reproduction and the body growth) cannot keep pace with the removal caused by fishery, and a further increase in exploitation level leads to a reduction in yield.

The fishing effort level which in the long term gives the highest yield is indicated by F(MSY) and the corresponding yield is indicated by "MSY", which stands for "Maximum Sustainable Yield". The phrase "in the long term" is used because one may achieve a high yield in one year by suddenly increasing the effort, but then meager years will follow, because the resource has been fished down. Normally, we are not aiming at such single years with maximum yield, but at a fishing strategy which gives the highest steady yield year after year.

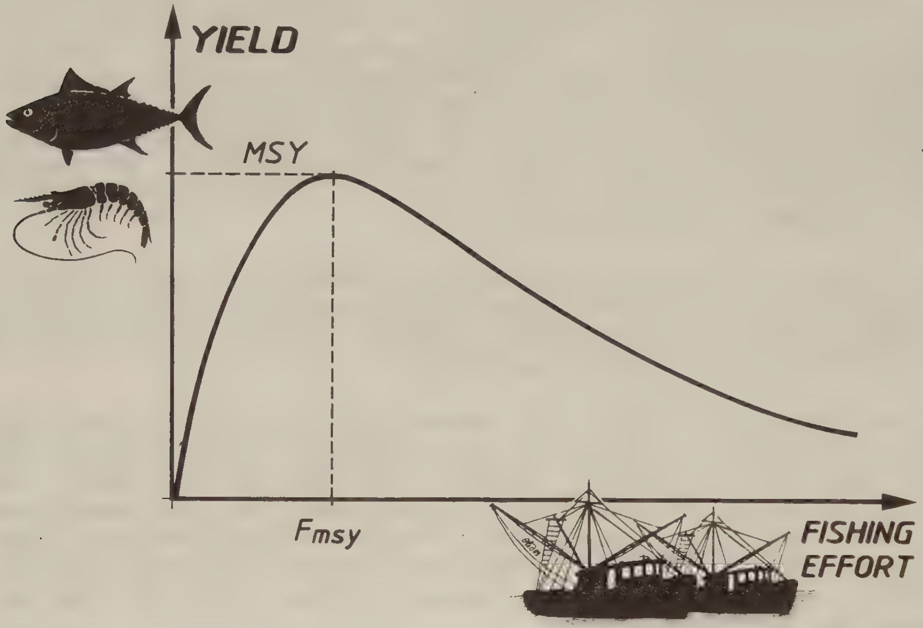


Fig. 1.1.1 **The basic objective of fish stock assessment**

1.2 The stock concept

When describing the dynamics of an exploited aquatic resource a fundamental concept is that of the "stock". By a "stock" we mean a group of organisms of one "species", having the same "stock parameters", and inhabiting a particular geographical area. The "stock parameters" are indicators of mortality and physiological features, for example, body growth.

A stock is a subset of a "species", which is generally considered as the basic taxonomic unit. A prerequisite for the identification of stocks is the ability to distinguish between different species. Because of the great number of different, but often similar species observed in tropical fisheries their identification can be problematic. The fisheries biologist, however, must master the technique of species identification if any meaningful fish stock assessment is to come out of the data collected. An aid to solve problems in species identification is provided by the so-called "FAO species identification sheets for fisheries purposes", (Fischer, 1978; Fischer and Bianchi, 1984; Fischer, Bianchi and Scott, 1984; Fischer and Hureau, 1986; Fischer, Schneider and Bauchot, 1987; Fischer and Whitehead, 1974) and in the "FAO species catalogues" (Allen, 1985; Colette and Nauen, 1983; Compagno, 1984 and 1984a; Holthuis, 1980; Roper, Sweeney and Nauen, 1984; Whitehead, 1985 and Whitehead, Nelson and Wongratana, 1988).

To the definition given above we can add that "stocks" are discrete groups of animals which show little mixing with the adjacent groups. One essential feature is that the population parameters remain constant over the distribution area of the stock, so that we can operate with the "stock parameters".

This definition may be too superficial for the taste of many biologists, and in the following a few more aspects of the stock concept are mentioned. A group of animals for which the geographical limits can be defined and therefore also the fisheries (the fishing fleets) which exploit this group of fish may be considered "a stock" in terms of fish stock assessment. Such a group of animals should belong to the same race within the species, i.e. be a group of animals which share a common gene pool. For species showing little migratory behaviour (mainly demersal species) it is easier to identify a stock than for highly migratory species, such as tunas.

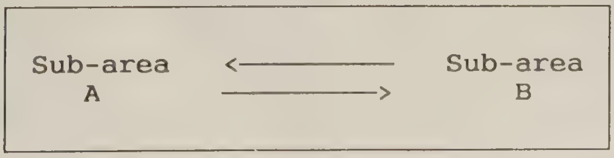
Cushing (1968) defines a fish stock as one that has a single spawning ground to which the adults return year after year. Larkin (1972) defines a stock as "a population of organisms which, sharing a common gene pool, is sufficiently discrete to warrant consideration as a self-perpetuating system which can be managed", while Ihssen et al. (1981) define a stock as "an intraspecific group of randomly mating individuals with temporal or spatial integrity".

Perhaps, the most suitable definition in the context of fish stock assessment was given by Gulland (1983) who stated that for fisheries management purposes the definition of a "unit stock" is an operational matter, i.e. a group of organisms can be treated as a stock if possible differences within the group and interchanges with other groups can be ignored without making the conclusions reached invalid.

A definition of the term "stock" acceptable to everyone with an interest in intraspecific groupings may be unattainable. For reviews of the stock concept, see Booke (1981), Ihssen et al. (1981) and MacLean and Evans (1981).

Fish stock assessments should be made for each stock separately. The results may (or may not) subsequently be pooled into an assessment of a multispecies fishery. Therefore, the input data must be available for each species considered. The stock concept is closely related to the concept of "population parameters", of which the so-called "growth parameters" are an example. The latter are numerical values in an equation by which we can predict the body size of a fish when it reaches a certain age. Other population parameters are the "mortality coefficients", which reflect the rate at which the animals die, i.e. the number of deaths per time unit. The mortality rates considered in this manual are the "fishing mortality", which reflects the deaths created by fishing and the "natural mortality", which accounts for all other causes of death (predation, disease etc.).

An essential characteristic of a stock is that its population parameters remain constant throughout its area of distribution. Let us, as an example, partition that area into two parts, sub-areas A and B:



The population parameters must be the same in sub-areas A and B, or in other words:

1) The animals in sub-area A must have the same body growth rate as the animals in sub-area B

2) The animals in sub-area A must have the same probability of death as the animals in sub-area B

If, for example, fishing takes place only in sub-area A, it is assumed that each individual fish in the stock has the same probability of being encountered in sub-area A and thereby also that it has the same probability of being caught. The individuals are supposed to move freely between the two sub-areas.

In order to determine whether a species forms one or more distinct stocks, we should examine its spawning areas, population parameters and morphological and genetic characteristics. We should also compare the fishing patterns in various areas and carry out tagging studies. The process is complicated, and often it is not possible with the knowledge in hand to determine whether there are several stocks of that species or not. There are two main reasons for failing to define a stock properly:

1) The full distribution area of the stock is not covered, so that only a part of the stock is considered, or the opposite

2) Several independent stocks are lumped together, for example because their areas of distribution overlap

Several countries may exploit the same stock. This is the case for many migratory stocks, e.g. tunas. It sometimes happens that a country assesses such a "shared" stock as if it were a national stock only exploited by that country. On the other hand, a fishery of a single country may exploit several independent stocks. Coral reef fish stocks may fall in this category.

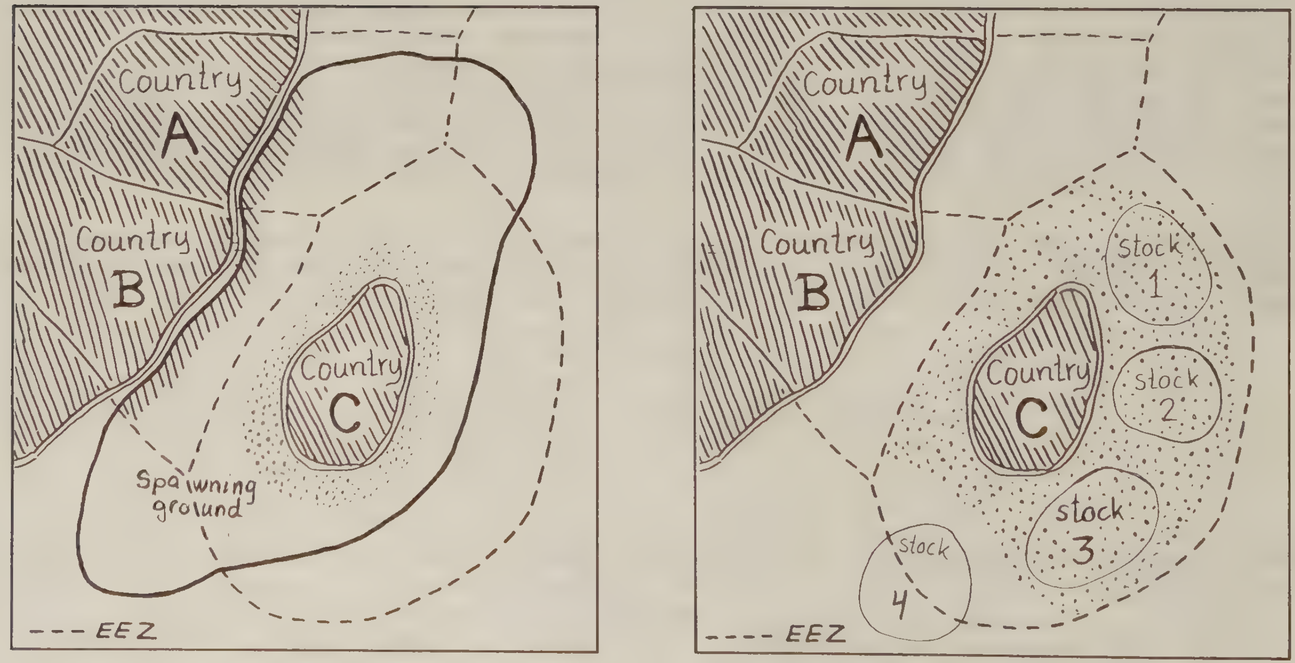


Fig. 1.2.1 Distributions of stocks related to management problems

Fig. 1.2.1 illustrates those two cases. In part I, we consider a fish stock, of which the distribution area is indicated by a full line. It is exploited by three countries, A, B and C, and we look at the stock definition from the point of view of the island country C. The broken lines show the EEZ (Exclusive Economic Zone) of each country, i.e. the national jurisdiction over fisheries. The dotted area indicates the fishing area of country C and the hatched areas those of countries A and B. From this it can be seen that if country C would base its assessment on the assumption that the unit stock is limited to its own fishing area, thus ignoring the fisheries of countries A and B, it is likely to draw wrong conclusions. If, for example, country A and B have intensive fisheries on the stock in question so that the stock is over-exploited (i.e. a reduction of the fishing intensity would increase the yield) there is little country C can do on its own to improve the situation. From the assessment based on the assumption of a stock limited to country C's waters, country C may conclude that the stock is over-exploited and it may introduce management measures to reduce fishing. However, the expected effect of the management measures will not materialize, if countries A and B do not follow country C.

Part II of Fig. 1.2.1 illustrates the case where one fishery exploits several stocks. In this case the assessment becomes that of the average stock, since it will be impossible to separate the catches by stocks. If the fishing effort expended is similar for each stock, the result of the assessment should come out correctly. However, there may also be difficulties in this case. Suppose that the three stocks currently fished (1, 2 and 3) are heavily overfished, and that the fishery is expanded to include the unexploited stock 4. In that case the average catch rate will increase and this might lead to wrong conclusions regarding the status of stocks 1, 2 and 3.

Nearly all exploited marine organisms undertake migrations, for example to their spawning grounds. A basic key to an understanding of stock structures is the knowledge of migration routes. This can be obtained from tagging experiments, but also from data and information provided by the commercial fisheries. Often the fishermen know where the spawning grounds are and they know where the high concentrations of fish found are at different times of the year.

Some general conclusions may be drawn from the above. Firstly it is usually safer to assume that species in neighbouring fishing areas form one unit stock than to consider each separate fishery to exploit its own unit stock. Further, it is evident that proper assessments can only be carried out when the biology of the species, including its migrations, spawning habits etc., is fully understood. Fish stocks are not bound by human geographical limits and this means that proper assessments can only be made when such limits can be ignored through interstate or international cooperation.

1.3 Models

A description of a fishery consists of three basic elements:

1) the input (the fishing effort, e.g. the number of fishing days)

2) the output (the fish landed) and

3) the processes which link input and output (the biological processes and the fishing operations)

Fish stock assessment aims at describing those processes, the link between input and output and the tools used for that are called "models". A model is a simplified description of the links between input data and output data. It consists of a series of instructions on how to perform calculations and it is constructed on the basis of what we can observe or measure, such as for example fishing effort and landings.

The actual processes which lead from a certain number of days fishing with a certain number of boats to a certain number of fish being landed are extremely complicated. However, the basic principles are usually well understood, so that by processing the input data by aid of models we can predict the output.

INPUT → PROCESSES → OUTPUT
observation model observation

A model is a good one if it can predict the output with a reasonable precision. However, since it is a simplification of reality it will rarely (and only by chance) be exact.

The instructions for the calculations that make up the model are given in the form of mathematical equations. These are composed of three elements: "variables", "parameters" and "operators". For example, the mathematical equation:

$$Y = 2.5 + 3 * X$$

has the variables Y and X, the parameters 2.5 and 3 and the operator "+". The equation is used to predict the value of Y for some value of X.

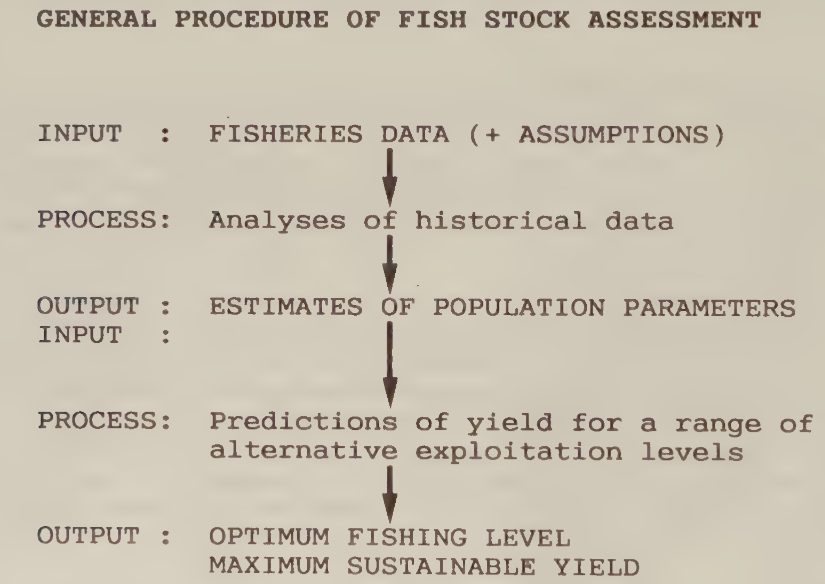

GENERAL PROCEDURE OF FISH STOCK ASSESSMENT
INPUT : FISHERIES DATA (+ ASSUMPTIONS)
PROCESS: Analyses of historical data
OUTPUT : ESTIMATES OF POPULATION PARAMETERS
INPUT :
PROCESS: Predictions of yield for a range of alternative exploitation levels
OUTPUT : OPTIMUM FISHING LEVEL
MAXIMUM SUSTAINABLE YIELD

Fig. 1.3.1 General flowchart for fish stock assessment

Fish stock assessment involves five basic steps as illustrated in Fig. 1.3.1. The first step is to collect data on the fishery, the INPUT to the assessment, which often have to be supplemented by assumptions or qualified guesses. Then we process the data by applying a model to estimate the population parameters, the OUTPUT from the processing of "the historical data". (The term "historical" is used to distinguish it from the subsequent process, the prediction of future yield). This prediction is based on the previous OUTPUT (= INPUT) and on a model, and the prediction is repeated for a series of alternative options. (Such options could be, for example, a fishing effort reduction of 10%, 20% and 30%, no change in fishing effort or a fishing effort increase of 10%, 20% and 30%). Among the alternative assumptions the best one is eventually selected as the final OUTPUT. The original INPUT data may be research survey data, data from samples drawn from the commercial fisheries or a combination of both.

Two main groups of fish stock assessment models are covered in this manual: "holistic models" and "analytical models". The simple holistic models use fewer population parameters than the analytical models. They consider a fish stock as a homogeneous biomass and do not take into account, for example, the length or age structure of the stock. The analytical models are based on a more detailed description of the stock and they are more demanding in terms of quality and quantity of the input data. On the other hand, as a compensation, they are also believed to give more reliable predictions.

The type of model to be used depends on the quality and quantity of input data. If data are available for an advanced analytical model then such a model should be used, while the simple models should be reserved for situations when data are limited. We are often in the situation where a complete set of input data for an analytical approach is not available, but where the available data exceed the demand of the simple models. As an alternative to using simple models in this case, the lacking input data can be replaced by assumptions or qualified guesswork. Often, the lacking parameter for a particular stock can be replaced by known parameters from another, similar stock.

1.3.1 Analytical models

A basic feature of analytical models as developed by, among others, Baranov (1914), Thompson and Bell (1934) and Beverton and Holt (1956), is that they require the age composition of catches to be known. For example, the number of one year old fish caught, the number of two year old fish caught, etc. may form the input data.

The basic ideas behind the analytical models may be expressed as follows:

1) If there are "too few old fish" the stock is overfished and the fishing pressure on the stock should be reduced

2) If there are "very many old fish" the stock is underfished and more fish should be caught in order to maximize the yield

(Some suggestions for more exact definitions of the term "overfishing" are given in Chapter 8).

The analytical models are "age-structured models" working with concepts such as mortality rates and individual body growth rates.

The basic concept in age structured models is that of a "cohort". To put it simply, a "cohort" of fish is a group of fish all of the same age belonging to the same stock. (We shall further elaborate on the definition of a cohort in Chapter 4). For example, a cohort of the threadfin bream, (*Nemipterus marginatus*) could be all the fish of that species that hatched from June to August in 1976 near Tanjung Pinang in the South China Sea. Suppose there were one million specimens in that cohort. After August 1976 the original one million fish would gradually decrease in number because of deaths due to either natural causes (predation, diseases, etc.) or fishing. However, while the number of survivors of the cohort decreases with time the average individual body length and body weight increase.

Fig. 1.3.2 shows an (hypothetical) example of the dynamics of a cohort, in the form of plots against age of the number of survivors (A), body length (B), body weight (C) and total biomass (D). Curve A shows the decay in the

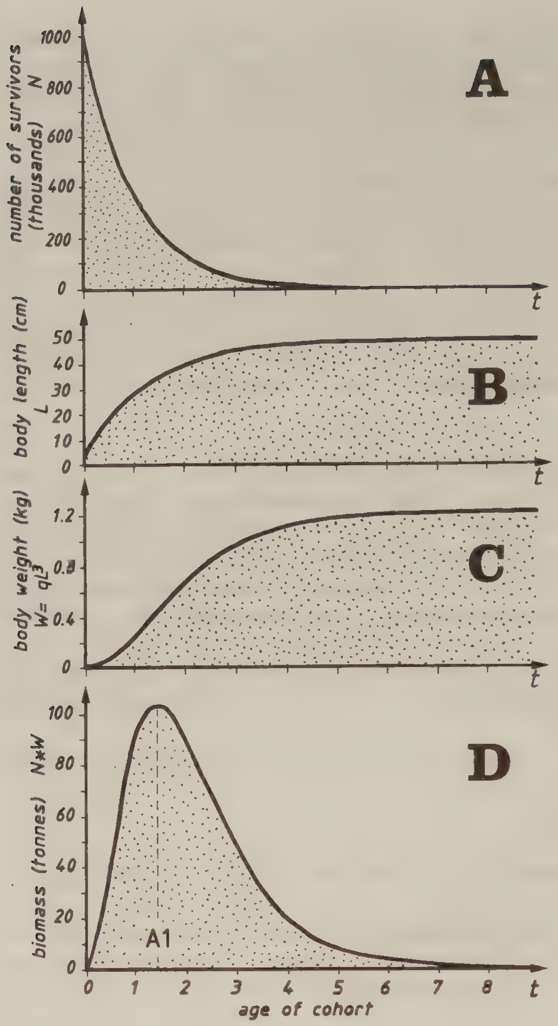


Fig. 1.3.2 The dynamics of a cohort

number of survivors as a function of the age of the cohort. Curve B shows how the average body length increases as the cohort grows older. Curve C shows the corresponding body weight, while curve D is a plot of the total biomass of the cohort, i.e. the number of survivors times the average body weight against the age of the cohort.

Note that curve D has a maximum (at age A1). Thus, to get the (hypothetical) maximum yield in weight from that cohort all fish should be caught exactly when the cohort has reached age A1. This, of course, is not possible in practice. However, you may say that the goal of fish stock assessment is to manage fisheries in such a way that catches come as close as possible to this theoretical maximum.

The implication is that the fish should neither be caught too young nor too old. If the fish are caught too young there is "growth overfishing" of the stock.

There are thus two major elements in describing the dynamics of a cohort:

1) The average body growth in length and weight

2) The death process

Both elements will be dealt with in great details in Chapters 3 and 4 respectively.

1.3.2 Holistic models

In situations where data are limited, for example, when starting up the exploitation of an hitherto unexploited resource, or in cases of limited capability of sampling, one may not have input data of the quality and in the quantity required for an analytical model. One solution would be to start up the collection of the data types required for the analytical approach and then wait until a sufficient amount is available. This approach is, of course, recommendable, because it solves the problem in the long run, but that may take years, while often advice on an exploitation or development strategy may be needed now. The approach taken in this manual is that no matter which type of data you have, there is always some information to be extracted from it, and that advice based on an analysis of a limited data set is usually better than complete guesswork.

In order to cover such data-limited situations, some simple holistic, less data demanding methods have been included in the manual. These methods disregard many of the details of the analytical models. They do not use age or length structures in the description of the stocks, but consider a stock as a homogeneous biomass.

Two types of simple methods are presented, namely the "swept area method" (in Chapter 13) and the "surplus production model" (in Chapter 9).

The swept area method is based on research trawl survey catches per unit of area. From the densities of fish observed (the weight of the fish caught in the area swept by the trawl) we obtain an estimate of the biomass in the sea from which an estimate of the MSY (Maximum Sustainable Yield) is obtained. This method is rather imprecise and it predicts only the order of magnitude of MSY.

The surplus production methods use catch per unit of effort (for example kg of fish caught per hour trawling) as input. The data usually represent a time series of years and usually stem from sampling the commercial fishery. The models are based on the assumption that the biomass of fish in the sea is proportional to the catch per unit of effort as shown in Fig. 1.3.3. An estimate of the yield is derived from the biomass.

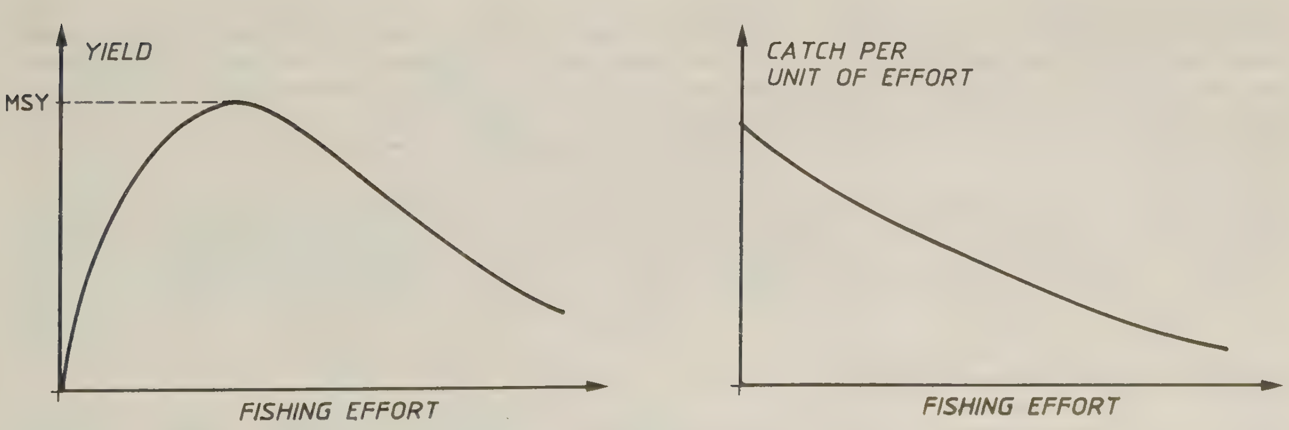


Fig. 1.3.3 Surplus production model

1.4 Assessment of stocks in tropical waters

The literature on fishery biology dealing with species in temperate zones is extensive compared to that on tropical fisheries. The major part of the literature on tropical fish stock assessment was published recently. As will appear in the following chapters of the manual, this can partly be attributed to the fact that tropical resources are somewhat more complex than those of temperate waters.

The present manual has the word "tropical" in its title. Although the methods described in the manual resemble those used in temperate waters, there are special features which justify the use of the word "tropical". Perhaps the most conspicuous difference between fish stock assessment in tropical waters and temperate waters lies in the nature of the basic input data rather than in the models.

For the analytical models we need the number of fish caught by each age group as input. In temperate waters stock assessment methods used are heavily dependent on the fortunate fact that ages of fish can be readily determined by "ageing" them. Ageing is most often done by counting rings in hard parts of the fish body, such as ear-bones (otoliths) or scales. The so-called year-rings are formed through a daily addition (daily-ring) to the size of the scale or otolith. The chemical composition and thereby the transparency of the addition depends (among other things) on the amount of food available and is therefore seasonal. The difference in deposits made in the winter and in the summer can be detected and one year-ring, composed of a summer and a winter part, can be distinguished from the next. Moreover, temperate fish species usually spawn once per year in a relatively short time-span, which makes it easy to distinguish year-classes or cohorts.

Also in tropical fish material is added daily to hard parts, which can be distinguished as daily growth rings. However, the lack of a strong seasonality makes the distinction of seasonal rings and therefore also of year-rings problematic for many tropical species. Moreover, the same absence of strong seasons results in less distinct spawning periods for most species. Many tropical species spawn at least twice per year and often over long periods. Fortunately, due to periodic changes in winds (monsoons) and shifts in oceanographic conditions (upwelling) in many tropical areas, a certain level of seasonality can still be detected. This seasonality may be reflected in the spawning patterns and growth of tropical fish species albeit less pronounced and much more difficult to detect than in temperate waters. These seasonal differences make it possible to detect also in tropical species the existence of different cohorts (often two per year), through the analyses of length-frequency samples.

In recent years, techniques have been developed to read daily rings in the otoliths of many fish species. This has enabled the development of age reading on tropical species, in particular of fish with short life-spans, or young fish. These techniques are still very time consuming and will be difficult to apply on a routine basis. They may however, serve to validate the results obtained from the analyses of length frequencies.

A further complication of tropical fish stock assessment vis-à-vis that in temperate waters is that the number of species caught in some important gears, in particular the bottom trawl, is very high. This does not only affect sampling and data collection procedures, it also makes it more difficult to apply the models. For a further discussion of differences and similarities between exploited stocks in arctic, temperate and tropical waters, see Ursin (1984).

The above mentioned differences can easily explain the slow rate of development of fish stock assessment in the tropics compared to that in temperate areas. The present manual works with methods which are the length-

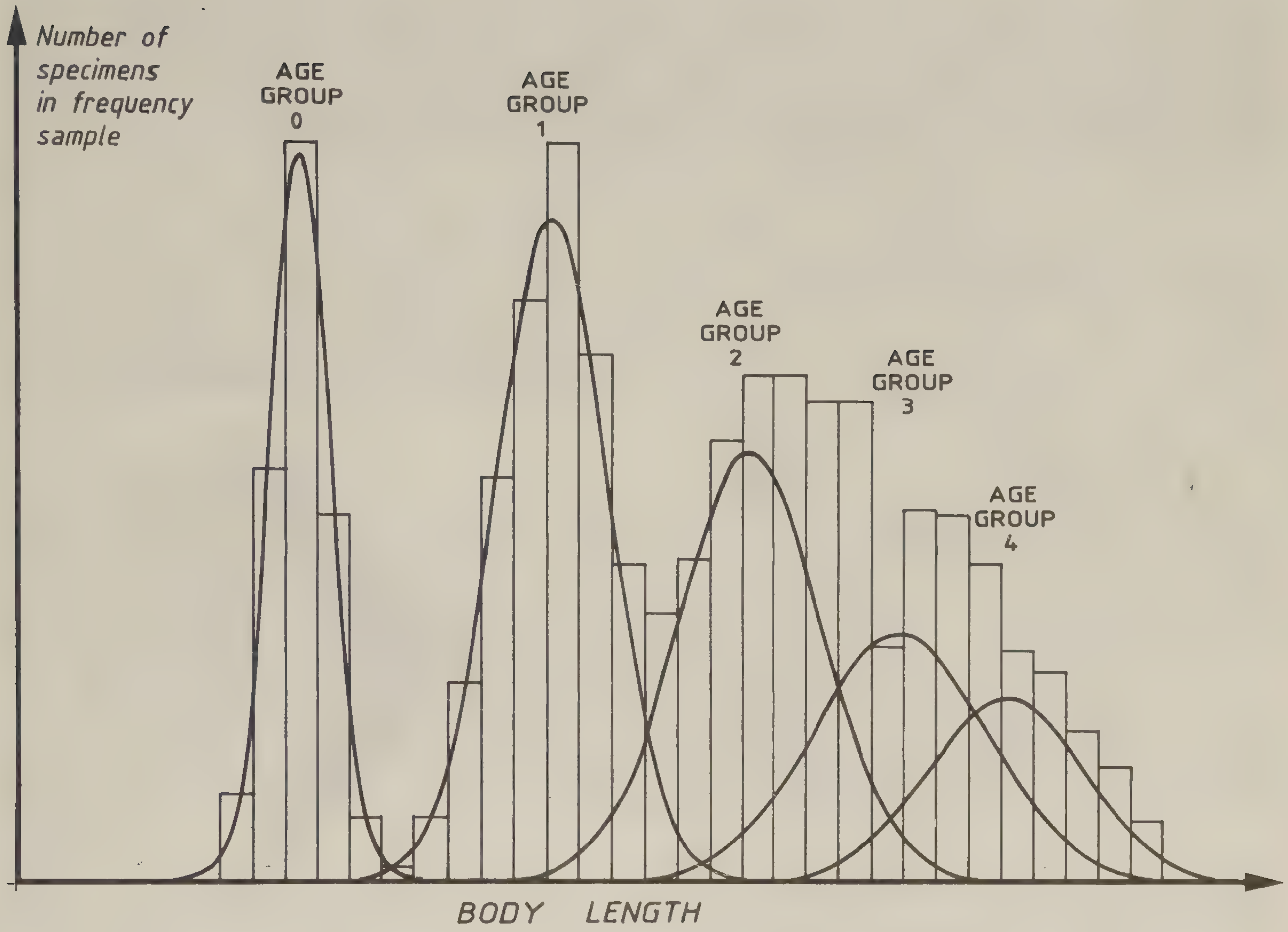


Fig. 1.4.1 Length frequency sample resolved into age groups

based parallels to the traditional age-based methods of temperate waters. Clearly, there is a relationship between age and length, and if the relationship is known we can convert length frequencies into age frequencies. Fig. 1.4.1 shows a resolution of a length frequency sample into age groups (cohorts). There are several techniques available for the separation of length groups and conversion into age groups, most of which are computerized. Several of these are discussed in the manual and one of them, the Bhattacharya method, is illustrated by examples and exercises. This method, although applicable in several computerized versions, can also be performed by using simply paper, pencil and a (scientific) pocket calculator.

In this manual, when explaining the theory behind the various methods we often start with the age based version, because it is easier to explain and consequently also easier to understand. The next step is then to convert the age based method into a length based method by using the relationship between age and length.

1.5 Definitions of body length

In the present context, "body length" means the average body length of a cohort. Individual fish are not considered in the models. When talking about "the length of an animal" in connection with a model it is always tacitly assumed that it is the "average length of the animals of a cohort". The estimate of average length, however, is derived from averaging the length measurements of individual specimens. The actual measure used for body length is not important as far as the theory behind the growth model is concerned. It is common practice to use the "total length" measured to the "nearest unit below" unless anatomical details make it not practicable (see Fig. 1.5.1). "Fork length" may be used for fish with stiff caudal fins (tunas) or special fin shapes (Nemipteridae). The "standard length" is not recommended for length frequency sampling.

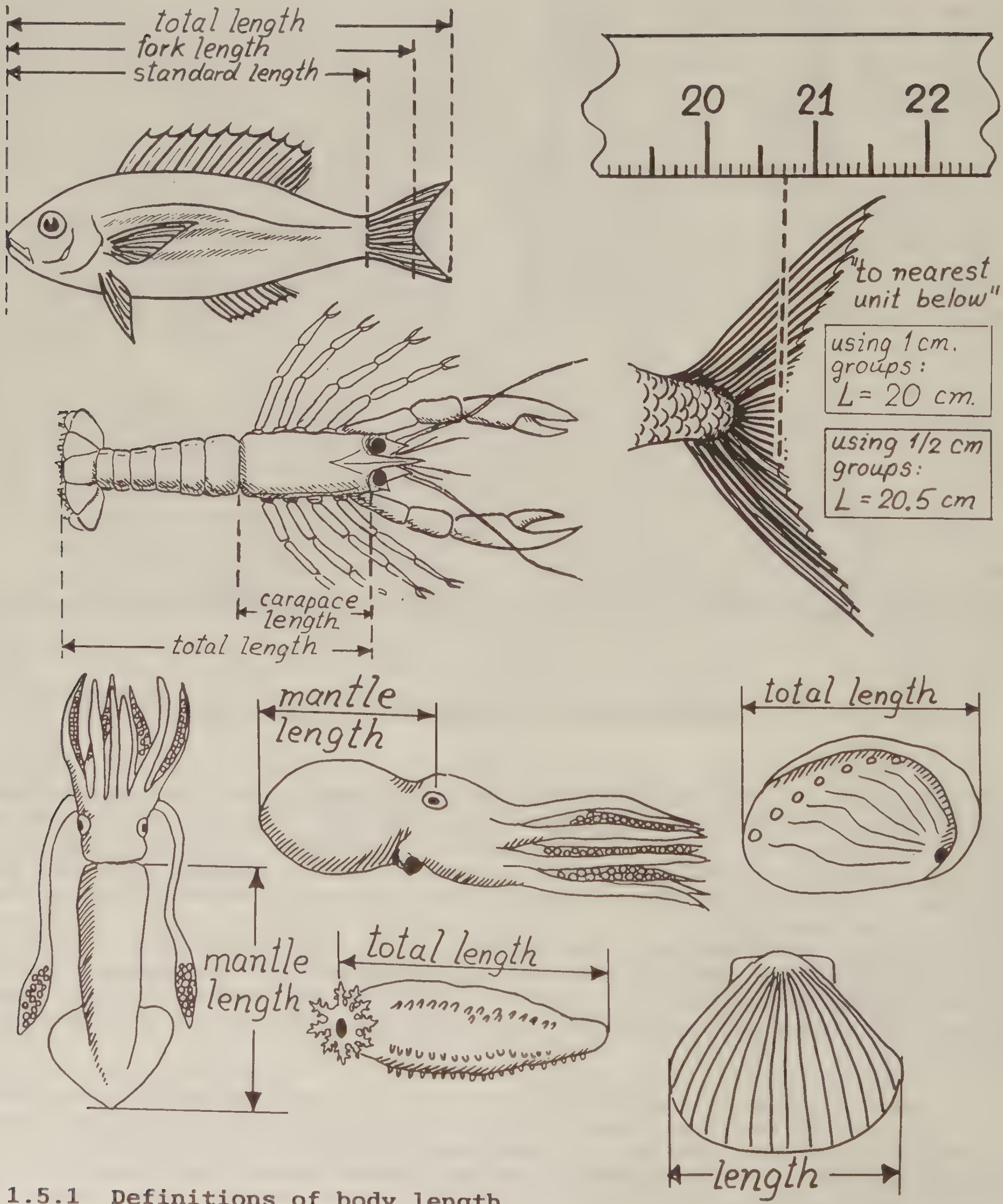

total length
fork length
standard length
20
21
22
"to nearest unit below"
using 1 cm. groups: L = 20 cm.
using 1/2 cm groups: L = 20.5 cm
carapace length
total length
mantle length
total length
mantle length
total length
length

Fig. 1.5.1 Definitions of body length

The most accurate measure for shrimps and lobsters is the "carapace length" (see Fig. 1.5.1). However, in may cases either total length or tail length will have been used. In such cases it is necessary to establish the relationship between the various measurements.

A really important thing is to specify exactly what kind of length measurement has beed used, as one may otherwise run into difficulties when comparing results with those of other investigations.

Other examples given in Fig. 1.5.1 are squid, octopus, abalone, sea cucumber and scallops. For animals with a hard shell or skeleton it is not a problem to define a suitable length measure (fish, crustaceans and molluscs with shell). Also molluscs with a relatively constant body form (e.g. squid) create no major problem, but animals with a plastic body (e.g. octopus, sea cucumber or jellyfish) are problematic. It may in certain cases be preferable to work with body weight rather than length, as the former is obviously measureable with greater accuracy.

It is easy to transform one type of length measurement into another type for a single individual. In cases where a sample is grouped into length classes it is more cumbersome to change from one measurement into another as far as the computational aspects are concerned. One simple way of doing it by microcomputer is given in Sparre (1987).

1.6 Age and recruitment

When working with analytical models we need to define the concept of "age". As was said above in connection with body length, we do not operate at the individual specimen's level, so "age" means the average age of a cohort. To define age we must start with a definition of "birthday". The obvious biological definition of the day of birth is the day the larva hatches from the egg. We say that a newly hatched fish has age zero.

In the first part of their life the larvae (or juveniles) are usually little influenced by the fishery. We say that the fish is then in the unexploited phase of life. Because we are interested in the exploited phase of its life the unexploited phase is not important in the present context.

Let Tr be the youngest age at which the fish may be vulnerable to fishing gears. A fish of age Tr is called a "recruit". By "recruitment" we mean the number of recruits, i.e. the number of fish that have attained age Tr during a "recruitment season". The "recruitment intensity" is the number of recruits per time unit. The recruitment pattern of a temperate species could be the as shown in Fig. 1.6.1A, where each line represents the recruitment intensity in one week. In most tropical fish stocks recruitment continues (more or less) all year round, but with seasonal oscillations, for example where monsoons occur (Pauly and Navaluna, 1983) (see Fig. 1.6.1B).

Let us tentatively define the recruitment season of a tropical fish stock by the dates (fractions of the year) tr1 and tr2 which correspond to the dates of minimum recruitment (see Fig. 1.6.1B). With 0 <= tr1 < tr2 <= 1.0 we define the "spring cohort" as the fish recruited from time tr1 to tr2 and the "autumn cohort" as the fish recruited from time tr2 to tr1. ("Spring" and "autumn" refer here to the northern hemisphere).

In general, the recruitment patterns of tropical fish stocks are not very well understood at present. However, as will appear from the following chapters, the seasonality in recruitment is a very important prerequisite for the methods suggested.

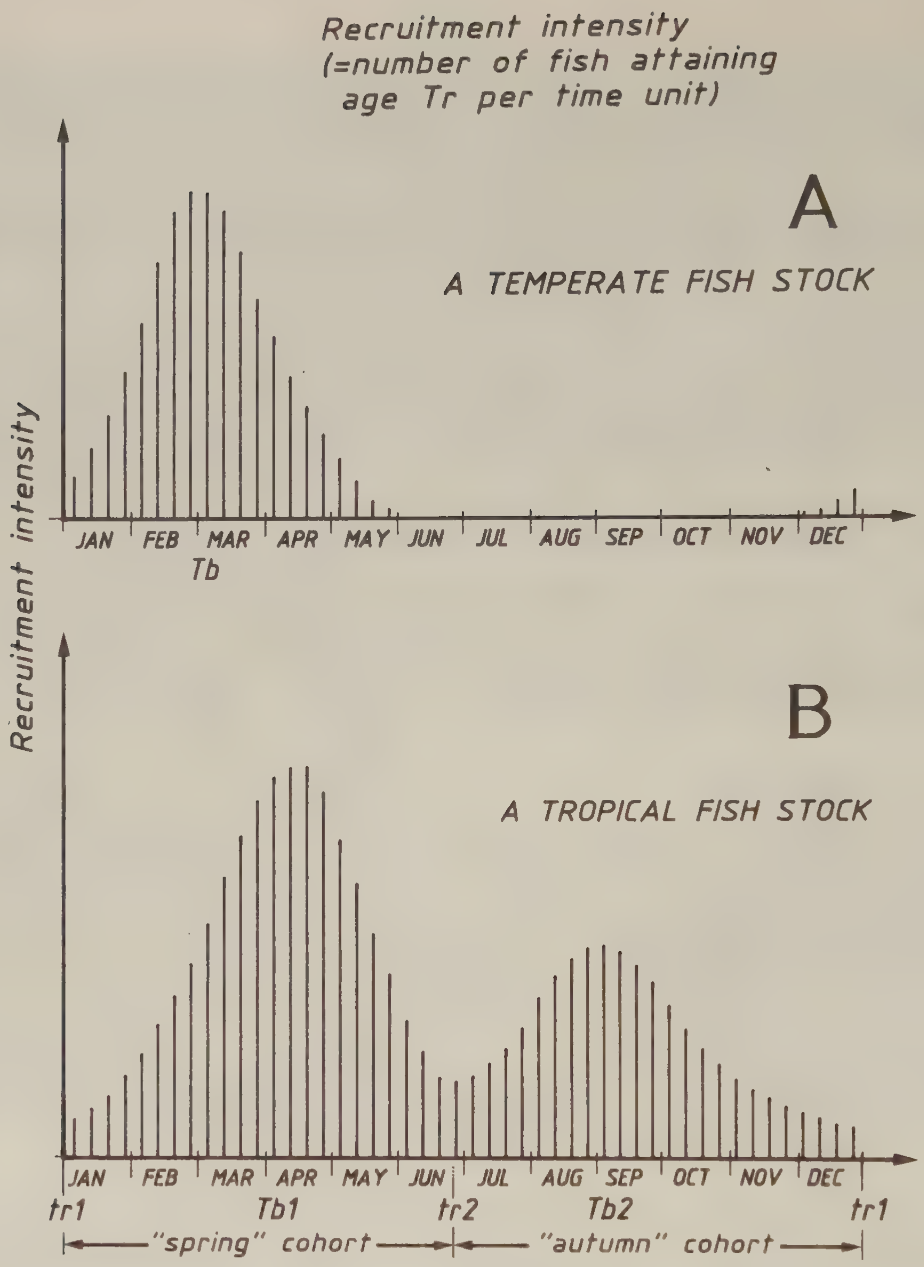


Fig. 1.6.1 Recruitment intensity during the year of typical temperate and tropical stocks

1.7 The underlying assumption of random samples

All the basic versions of the methods dealt with in the manual assume the input data to be derived from "random samples". A sample of fish, for example a length frequency sample representing the stock, is a random sample if any fish in the entire stock has the same probability of being drawn.

Usually, it is difficult or even impossible to obtain pure random samples. If, for example, the juvenile fish are located in certain nursery areas, which do not coincide with the fishing grounds from which our samples originate, the juvenile fish will be underrepresented in the samples. A similar problem is created by the selectivity of fishing gears. Often the small fish are underrepresented because they escape through the meshes, whereas the larger fish are retained. Samples which are not random samples are called "biased samples".

A feature of fish behaviour which is believed to create the most serious bias is "migration". Almost all marine animals perform systematic movements. Pelagic fish such as mackerels, scads and tunas undertake long migrations between feeding grounds and spawning grounds. Most penaeid shrimps start their life cycle in the open sea and migrate to shallow waters (lagoons and mangroves) and when sexually mature they migrate back to the open sea for reproduction.

The implication of the migratory behaviour is that a large sea area must be covered in order to obtain random samples from the entire population. Often samples can be obtained only from the commercial fishery which concentrates on those grounds where the resources are easiest to catch in large quantities. Thus, we are often in the situation that random samples of the population are not available. This bias must be accounted for in the analysis and the basic methods have to be modified to account for it. Some types of bias ae easier to deal with than others. Bias created by migration can only be handled properly when the migration routes are known. When they are not we have to make certain assumptions about them in order to get on with the analysis. There are many serious problems in connection with bias. A few suggestions on how to get around them are presented, but the manual also leaves a number of relevant questions open, either because the author does not know the answer or the method is so complicated that it falls outside the scope of this manual. Unfortunately, one often comes across cases in practice, which are so heavily influenced by bias that they cannot be handled by the methods described here (see Chapter 11).

1.8 The organization of the manual

The complexity of fish stock assessment is reflected in the contents table of this manual. The various elements have to be dealt with in chronological order and it has often been necessary to refer to earlier or later sections and chapters. In order to assist the reader (and teacher) a flowchart for fish stock assessment as presented in this manual is given in Fig. 1.8.1. The flowchart does not present the methods in the same sequence as they appear in the manual, but rather in the natural chronological order of a fish stock assessment. The numbers of the relevant chapters are given in brackets.

Before starting as a fish stock assessment worker there are a few general basic statistical techniques (for example linear regression analysis) one must master. These are dealt with in Chapter 2. They have been placed outside the proper flowchart, because the methods are general and applied in many other scientific fields. Chapter 2 contains only a small selection of statistical methods and only those which are needed to follow the text in the next chapters.

The flowchart is divided into two parts. Part A deals with analytical methods and Part B deals with holistic methods. As the sizes of the two parts indicate, the main emphasis has been placed on the analytical methods. Both approaches follow the same main lines, namely the set-up given in Fig. 1.3.1.

Part A, the analytical methods

The first row shows the input to the estimation of growth parameters (the parameters by which we can predict the length of an animal for a given age). Although collection of data comes first chronologically, it is not dealt with in the beginning of the manual, because one cannot deal with data collection in a meaningful way before the objectives of the sampling scheme have been defined. To define the objectives we need the models used for the analysis of historical data, therefore, the main text on data collection is deferred to Chapter 7. The assumptions indicated as input are not dealt with in a particular chapter.

Part A

ANALYTICAL METHODS

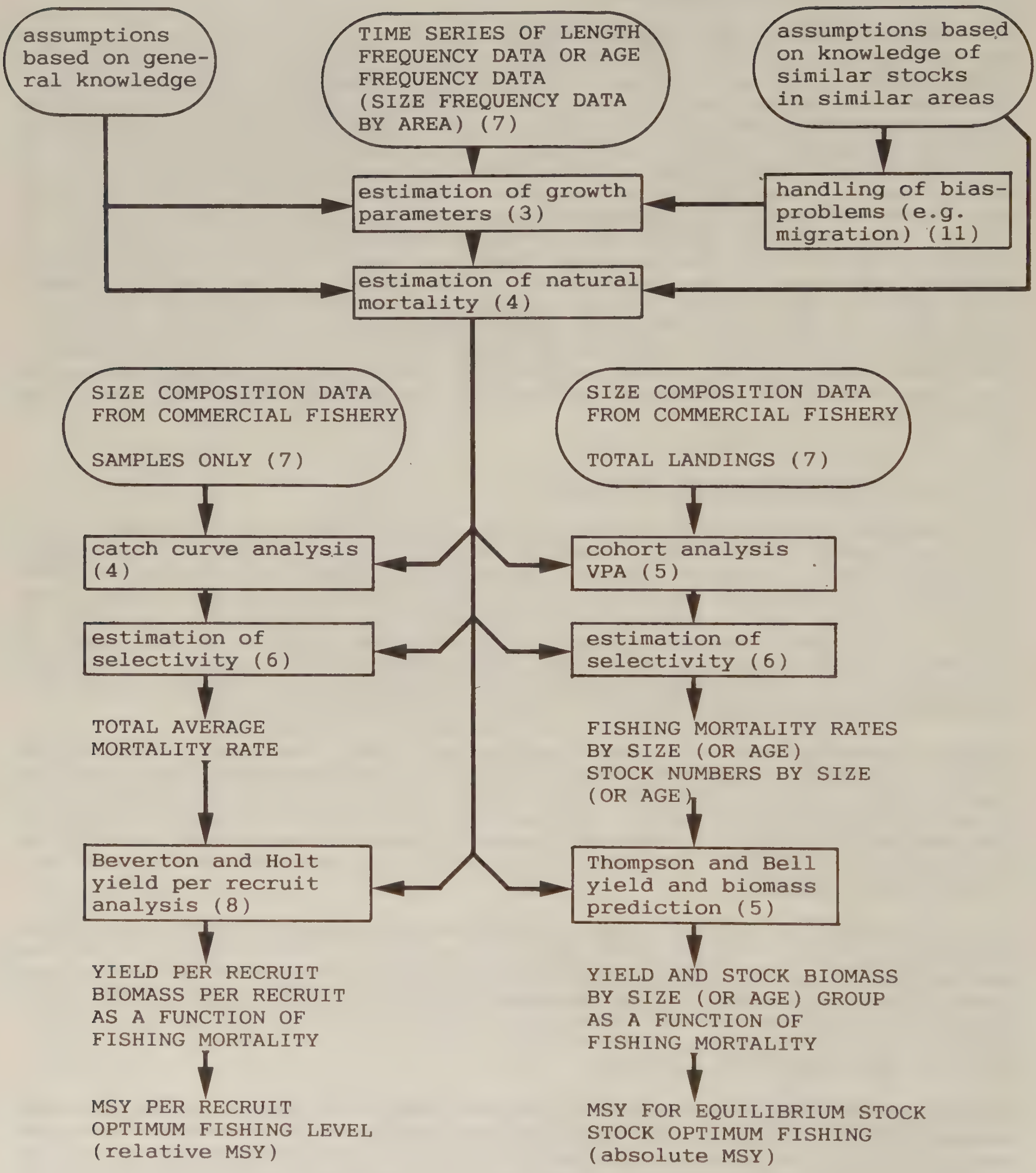
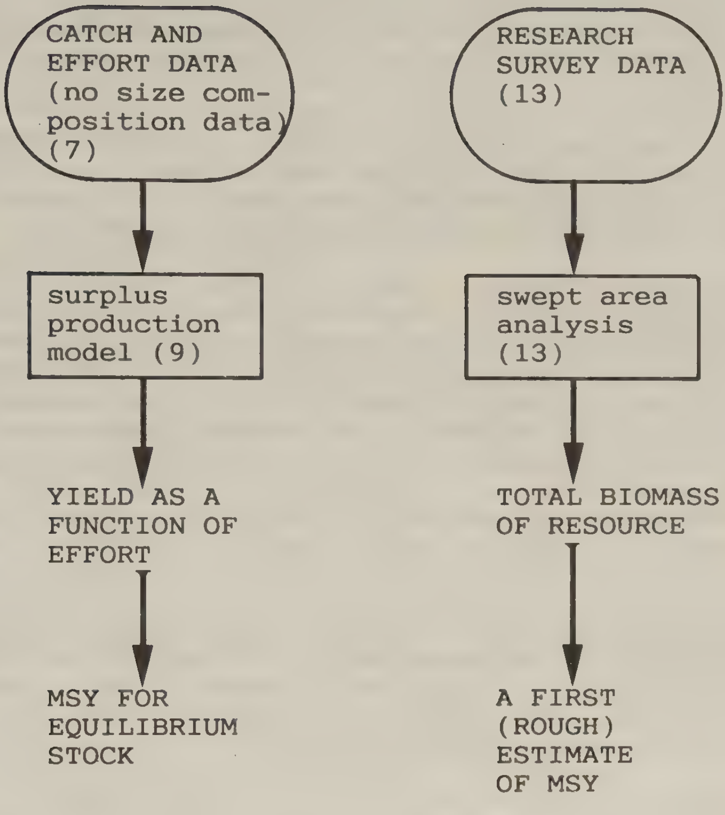

assumptions based on general knowledge
TIME SERIES OF LENGTH FREQUENCY DATA OR AGE FREQUENCY DATA (SIZE FREQUENCY DATA BY AREA) (7)
assumptions based on knowledge of similar stocks in similar areas
estimation of growth parameters (3)
handling of bias-problems (e.g. migration) (11)
estimation of natural mortality (4)
SIZE COMPOSITION DATA FROM COMMERCIAL FISHERY SAMPLES ONLY (7)
SIZE COMPOSITION DATA FROM COMMERCIAL FISHERY TOTAL LANDINGS (7)
catch curve analysis (4)
cohort analysis VPA (5)
estimation of selectivity (6)
estimation of selectivity (6)
TOTAL AVERAGE MORTALITY RATE
FISHING MORTALITY RATES BY SIZE (OR AGE) STOCK NUMBERS BY SIZE (OR AGE)
Beverton and Holt yield per recruit analysis (8)
Thompson and Bell yield and biomass prediction (5)
YIELD PER RECRUIT BIOMASS PER RECRUIT AS A FUNCTION OF FISHING MORTALITY
YIELD AND STOCK BIOMASS BY SIZE (OR AGE) GROUP AS A FUNCTION OF FISHING MORTALITY
MSY PER RECRUIT OPTIMUM FISHING LEVEL (relative MSY)
MSY FOR EQUILIBRIUM STOCK STOCK OPTIMUM FISHING (absolute MSY)

Fig. 1.8.1 The organization of the manual

Part B SIMPLE HOLISTIC METHODS


CATCH AND EFFORT DATA (no size composition data) (7)
RESEARCH SURVEY DATA (13)
surplus production model (9)
swept area analysis (13)
YIELD AS A FUNCTION OF EFFORT
TOTAL BIOMASS OF RESOURCE
MSY FOR EQUILIBRIUM STOCK
A FIRST (ROUGH) ESTIMATE OF MSY

Basic methods of bio-statistics (2)

The following symbols have been used:

⬭ : Input data or assumptions

▭ : Processing of data using a model

text not framed CAPITAL LETTERS : Output from a data processing stage; an output from one processing stage acts as input for the subsequent processing stage

(n) : Figures in brackets refer to the Chapter with the relevant theory

→ : Flow of data

The theory behind the model for body growth and the estimation of growth parameters is dealt with in Chapter 3. Handling of bias problems is as mentioned above extremely complicated and only partly covered by the present manual. It has therefore been placed in Chapter 11 after the chapters dealing with the analytical methods in their basic versions. Although we should start the analysis with an evaluation of the bias, it has not been considered appropriate to start the manual with one of the most complicated subjects. Also the estimation or rather the qualified guessing on natural mortality is a tricky subject. It has been placed in Section 4.7.

The following chapters on analytical methods contain the theory for both age-based methods and length-based methods. The estimates of growth parameters are in fact only used for the length-based versions of the models, but in order to reduce the complexity of the flowchart, no distinction has been made between length-based and age-based methods.

After the growth part (Chapter 3) the flowchart branches. The two branches represent a grouping of methods according to their data requirements. Some analytical methods are based exclusively on samples from the commercial fishery, while the total catch is not known. In theory these methods could be used also on a single sample consisting of one bucket of fish sampled in the local fish market (although, of course, extensive sampling schemes are recommended). Such methods are called "catch curve methods".

Other methods are based on estimates of the total catch, i.e. on estimates of the total number landed in each length group from the entire stock. Such numbers are derived from length frequency samples by raising them to account for the entire catch using data on total landings. These methods are called "cohort analysis" or "Virtual Population Analysis" (VPA). Compared to the catch curve methods, they give more dependable estimates of the parameters and more reliable predictions of the future fishery. The sampling procedures to obtain the input data are discussed in Chapter 7 as mentioned above in connection with growth.

The general theory of the death process, the "exponential decay model", the "catch curve methods" and some other methods with similar or limited data requirements are dealt with in Chapter 4. From the catch curve analysis we obtain an estimate of the total mortality, which combined with growth parameters and natural mortality, allows us to arrive at some conclusions on the current state of the stock and its potential. This analysis is performed by the "Beverton and Holt yield per recruit model", which is presented in Chapter 8. The ultimate result is the "maximum sustainable yield per recruit" (MSY/R).

The methods based on the size composition of the total catch, the cohort analysis and the VPA are covered in Chapter 5. The results obtained from cohort analysis (or VPA) are estimates of absolute stock size and fishing mortality for each size group. These results are used for the prediction of stock biomass and yield levels using the "Thompson and Bell methods" which are also described in Chapter 5. The final result is an estimate of the (absolute) MSY.

In the discussion presented above only the regulation of fishing effort has been considered as an instrument for fisheries management. However, there are other instruments, one of which is the size selectivity of the fishing gear. For example, by using larger mesh sizes in the codend of a trawl, the mortality on young fish is reduced, which subsequently will increase the catch of older, larger and more valuable specimens. The effects of "gear selectivity" are discussed in Chapter 6.

In all the chapters mentioned so far, the theory has been presented in its most simple form, i.e. only one stock of fish and only one type of fishing boat are considered. In reality there hardly exists any fishery where only one species is caught by only one type of boat. In most cases we have to deal with multispecies catches made by a variety of different boats. In theory it does not create any major problems to extend the analytical models to deal with the multispecies/multifleet case, and this has been demonstrated for the Thompson and Bell model in Chapter 10. Some other aspects of multispecies assessment are also briefly discussed in this chapter. With this we have reached the end of Part A of the flowchart.

Part B, the holistic methods

This part of the flowchart is less complicated, because it describes simpler methods. The "surplus production models" are presented in Chapter 9 and the "swept area method" in Chapter 13.

Chapters 12 and 14 have not been included in the flowchart. Chapter 14 describes (briefly) a package of microcomputer programs, the so-called LFSA (Length based Fish Stock Assessment) package (Sparre, 1987), which matches the analytical models of this manual (Part A of the flowchart). Chapter 12 deals with the unsolved stock/recruitment relationship - the question of a possible relationship between recruitment and the size of the parent stock. The problem is discussed in essay form and Chapter 12 does not suggest any models for practical application.

1.9 Further reading

Since the scope of this manual is mainly limited to methods and their applications, it is advisable to supplement the knowledge derived from it, by reading additional textbooks and manuals on stock assessment and on the biology of the most important resources. Chapter 15 contains references to several recent publications, which will be of great use in providing a better insight in stock assessment problems, for example

1) The relationship between stock assessment and management (Pauly, 1979 and Gulland, 1988)

2) The biology and assessment of shrimps (Garcia and Le Reste, 1981; Gulland and Rothschild, 1984; Penn, 1984; Garcia, 1985; Rothlisberg, Hill and Staples, 1985 and the Australian Journal of Marine and Freshwater Research, 1987)

3) The biology and assessment of cephalopods and other invertebrates (Caddy, 1983, 1983a and 1989)

4) Resource mapping (Caddy and Garcia, 1986 and Butler et al., 1986)

5) Tunas (Sharp and Dizon, 1978; Kleiber, Argue and Kearney, 1983; I-ATTC, 1984 and Hunter et al., 1986)

6) Migration (Harden Jones, 1968, 1984 and Oxenford and Hunte, 1986)

The various methods presented in this manual were used either directly or through the LFSA and COMPLEAT ELEFAN packages at FAO/DANIDA follow-up training courses, where participants processed their own data. The results of the analyses and the input data were published (Venema, Christensen and Pauly, 1988), with the aim of providing additional examples of the application of stock assessment methods on tropical resources. In some cases the papers demonstrate clearly the limitations of the data sets and of the methods used, but there are also several examples of successful applications of methods hitherto seldom used in tropical areas.

2 BIOSTATISTICS

This chapter contains a brief description of some statistical methods in common use in tropical fishery biology and introduces the statistical notation adopted in the manual. It serves as a refresher and reference point, but is not intended as a textbook in its own right.

The amount of literature on statistical methods is staggering, so there is no problem in case you want to do further studies in biostatistics. Only two references are given here. The book "Biometry" by Sokal and Rohlf (1981) deals with the theory in a rather accessible way, while "Sampling techniques" by Cochran (1977) is perhaps a bit more complicated, but still recommended as an introduction. However, there are many other textbooks which may be equally useful.

2.1 Mean value and variance

Let us consider a sample of n fish all of one species caught in one trawl haul and let x(i) be the length of fish no. i, i = 1,2,...,n. The "mean length" (in general the "mean value") of the sample is defined:

$$\bar{x} = (x(1)+x(2)+...+x(n))/n = \frac{1}{n} \sum_{i=1}^{n} x(i) \qquad (2.1.1)$$

The two first columns of Table 2.1.1 show an example for n = 27.

The variance, which is a measure of the variability about the mean value is defined as follows:

$$s^2 = \frac{1}{n-1} \left((x(1)-\bar{x})^2 + (x(2)-\bar{x})^2 + ... + (x(n)-\bar{x})^2 \right)$$

$$= \frac{1}{n-1} \sum_{i=1}^{n} (x(i)-\bar{x})^2 \qquad (2.1.2)$$

Thus, the variance, s^2 is the sum of the squares of the deviations from the mean divided by the number, n, minus one. The third and fourth column of Table 2.1.1 illustrate the calculation of the variance. Note that if all fish in the sample had the same length this would equal the mean length and the variance would be zero. The sum of the deviations (not squared) is always zero. The larger the deviations from the mean value, the larger the variance will be. The two largest values of the square of the deviations from the mean in Table 2.1.1 occurred for the smallest and the largest observations.

The square root of the variance, s, is called the "standard deviation". Often one is interested in the variance relative to the size of the mean length, and for that purpose s is the relevant quantity as it has the same unit as the mean. This leads to the relative standard deviation $s/\bar{x}$, also called the "coefficient of variation".

When doing the calculations by hand it is easier to work with a rearranged form of Eq. 2.1.2, which is equivalent to

$$s^2 = \frac{1}{n-1} \left(\sum_{i=1}^{n} x(i)^2 - \frac{1}{n} \left(\sum_{i=1}^{n} x(i) \right)^2 \right) \qquad (2.1.3)$$

However, as most scientific pocket calculators contain an option for automatic calculation of mean and variance the calculations are here illustrated by Eq. 2.1.2, which is conceptually easier to understand.

For many purposes, e.g. for graphical representation, it is convenient to arrange the sample in the form of a "frequency table" by dividing the length range into a number of length intervals. The length range for the sample in Table 2.1.1 goes from 11.2 to 19.0 cm. With length groups of 1 cm we need nine length groups to cover the range. Using 10.5 as the lower limit of the first length interval, the intervals and the frequencies of lengths become those shown in the first four columns of Table 2.1.2, which is a so-called length frequency table.

Table 2.1.1 Mean value, variance and standard deviation of a length frequency sample

fish no.	length (cm)	deviation from mean	square of deviation from mean
i	x(i)	$x(i)-\bar{x}$	$(x(i)-\bar{x})^2$
1	14.2	-0.87	0.75
2	16.3	1.23	1.52
3	14.8	-0.27	0.07
4	13.2	-1.87	3.48
5	16.9	1.83	3.36
6	12.4	-2.67	7.11
7	14.3	-0.77	0.59
8	15.7	0.63	0.40
9	15.3	0.23	0.05
10	11.2 (min.)	-3.87	14.95
11	12.9	-2.17	4.69
12	13.5	-1.57	2.45
13	18.2	3.13	9.82
14	11.6	-3.47	12.02
15	18.5	3.43	11.79
16	16.3	1.23	1.52
17	15.5	0.43	0.19
18	15.8	0.73	0.54
19	13.2	-1.87	3.48
20	19.0 (max.)	3.93	15.47
21	12.0	-3.07	9.40
22	17.1	2.03	4.13
23	15.4	0.33	0.11
24	14.6	-0.47	0.22
25	14.0	-1.07	1.14
26	18.1	3.03	9.20
27 = n	16.8	1.73	3.00
Total	406.8	0.00	121.48
	$= \Sigma\, x(i)$		$= \Sigma(x(i)-\bar{x})^2$

mean length, $\bar{x}$: 406.8/27 = 15.07
variance, s^2 : 121.48/(27-1) = 4.67
standard deviation, s : $\sqrt{4.67}$ = 2.16
relative standard deviation, $s/\bar{x}$: 2.16/15.07 = 0.14
standard error, $s/\sqrt{n}$: $2.16/\sqrt{27}$ = 0.41

(The concept of standard error is introduced in Section 2.3)

Table 2.1.2 Mean and variance from a length frequency sample. (The sample is derived from Table 2.1.1 with a class length of 1 cm)

index j	interval (cm) $L(j)-L(j+1)$	mid-point (cm) $\bar{L}(j)$	frequency $F(j)$	$F(j)\bar{L}(j)$	$(\bar{L}(j)-\bar{x})$	$F(j)(\bar{L}(j)-\bar{x})^2$
1	10.5-11.5	11	1	11	-4.074	16.60
2	11.5-12.5	12	3	36	-3.074	28.35
3	12.5-13.5	13	3	39	-2.074	12.91
4	13.5-14.5	14	4	56	-1.074	4.61
5	14.5-15.5	15	4	60	-0.074	0.02
6	15.5-16.5	16	5	80	0.926	4.29
7	16.5-17.5	17	3	51	1.926	11.13
8	17.5-18.5	18	2	36	2.926	17.12
9	18.5-19.5	19	2	38	3.926	30.83
		total	27	407		125.86

mean length, $\bar{x}$: 15.07 = 407/27

variance, s^2 : 4.84 = 125.86/26

standard deviation, s : 2.20 = $\sqrt{4.84}$

relative std. dev., $s/\bar{x}$: 0.15 = 2.20/15.07

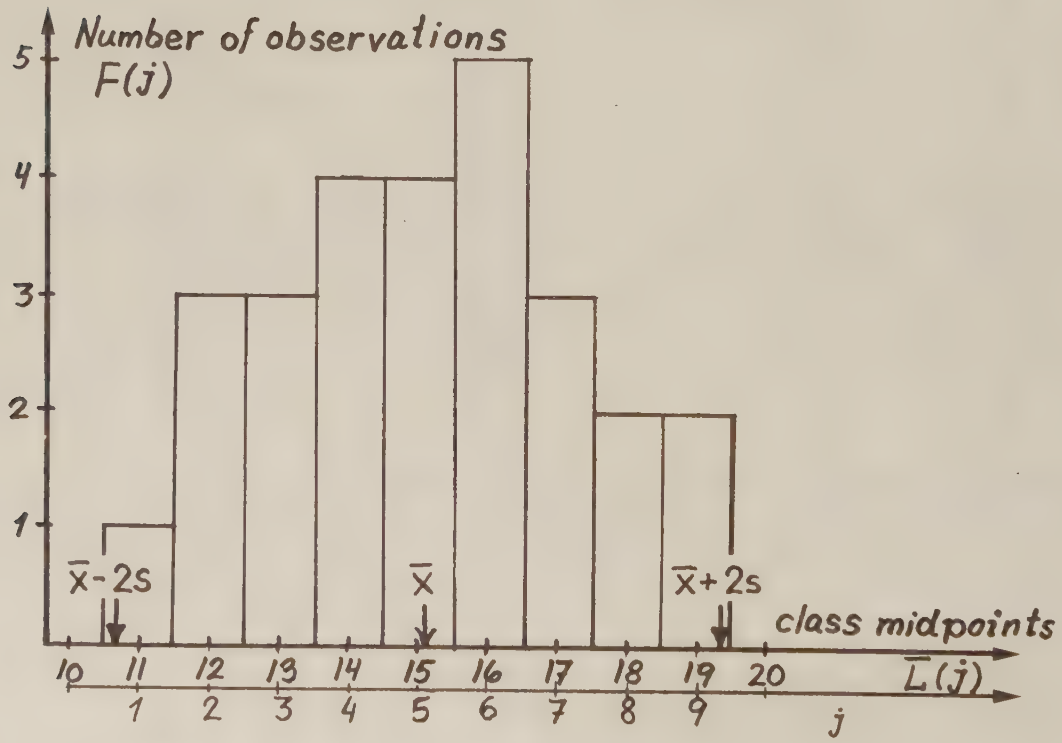


Figure 2.1.1 Length frequency diagram. Graphical representation of the frequency sample from Table 2.1.2

Let j be the index of a length group, and let the lower and upper class limit of length group no. j be denoted by respectively:

$$L(j) = L(1) + (j-1) * dL \quad \text{and} \quad L(j+1) = L(1) + j * dL,$$

$$\text{or} \quad L(j+1) = L(j) + dL$$

where dL is the interval size. A fish of length x(j) then belongs to length group j when

$$L(j) \;<=\; x(j) \;<\; L(j+1)$$

Let F(j) be the frequency of length group j, i.e. the number of fish observed in length group j. Let $\bar{L}(j) = L(j) + dL/2$ be the midpoint of length group no. j. The calculation of mean value and variance from a frequency table is then performed in the usual way using midpoints to represent the intervals:

$n = \sum_{j=i}^{m} F(j)$ is the total number of observations, where m is the number of length groups.

$\bar{x} = \frac{1}{n} \sum_{j=1}^{m} F(j)\, \bar{L}(j)$ is the mean value.

$s^2 = \frac{1}{n-1} \sum_{j=1}^{m} F(j)\, (\bar{L}(j) - \bar{x})^2$ is the variance.

The calculation procedure is shown in Table 2.1.2. The class midpoint $\bar{L}(j)$, and the square of the deviations from the mean are weighted by the number of fish in each class, i.e. the frequency, F(j). The results of Table 2.1.2 deviate slightly from those of Table 2.1.1 because a representation in cm groups produces less precise results than a representation in millimeter groups.

Figure 2.1.1 shows a graphical representation of the frequency sample. Note that all observations lie in the interval from

$$\bar{x} - 2s \text{ to } \bar{x} + 2s$$

For the so called normal distribution (discussed in the next section) we expect about 95% of the observations to be contained in that interval.

(See **Exercise(s)** in Chapter 17).

2.2 The normal distribution

Table 2.1.2 and Figure 2.1.1 show an example of a small set of length frequency data that approximately follows the so called "normal distribution". The mathematical expression for a normal distribution is:

$$Fc(x) = \frac{n\, dL}{s\sqrt{2\pi}} \exp\left[-(x - \bar{x})^2/(2\, s^2) \right] \qquad (2.2.1)$$

where Fc = "calculated frequency" or "theoretical frequency", n = number of observations, dL = interval size, s = standard deviation, $\bar{x}$ = mean length and π = 3.14159.

Using the values n = 27, dL = 1 cm, s = 2.20, $\bar{x}$ = 15.07 cm from Table 2.1.2 we get:

$$Fc(x) = \frac{27 * 1}{2.20 * \sqrt{2 * 3.14159}} \exp[- (x - 15.07)^2 / (2 * 4.84)] =$$

$$= 4.896 \exp(- (x - 15.07)^2/9.68)$$

The values of Fc for a number of different x-values are given in Table 2.2.1. Note that the notation is slightly modified as we now use the interval midpoint, x, as the argument in Fc instead of the interval index, j, as used for argument in F in Table 2.1.2.

Figure 2.2.1 shows the theoretical frequencies together with the bar diagram for F(j) from Figure 2.1.1. As can be seen, Fc(x) gives a fair fit to the observed length frequencies. This picture is often observed when recording length frequencies of fish originating from one cohort, i.e. fish of approximately the same age.

The normal distribution is observed in a great variety of different cases - hence the name. There are other types of probability distributions observed in fishery science. Examples are the "log-normal distribution", the "negative binomial distribution" and the "delta distribution". A conspicuous difference between these and the normal distribution is that they are skewed, whereas the normal distribution is symmetric. The delta distribution for example, is used to describe the probability distribution for the catch per hour by a trawl. It is composed of a log-normal distribution, which describes the distribution of the non-zero trawl catches and a special probability for zero catch (see Section 13.7, Fig. 13.7.2).

Table 2.2.1 Theoretical frequencies corresponding to Table 2.1.2

x	11	12	13	14	15	16	17	18	19
Fc(x)	0.88	1.85	3.14	4.35	4.89	4.48	3.33	2.02	0.99

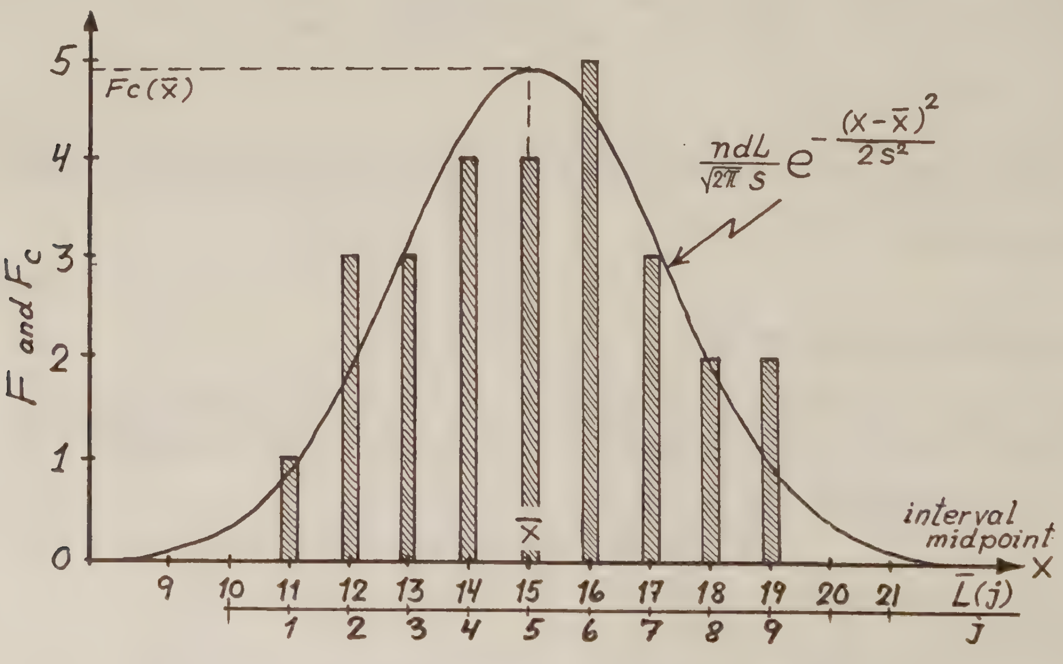


Figure 2.2.1 The theoretical frequency, Fc, (the normal distribution curve) and the observed frequencies, F, (bars)

Perhaps the most important feature about normal distributions has to do with mean values. If you take a number of, say 50, samples from a population and calculate the mean values for each sample you will get 50 different mean values. Thus, a mean value has a probability distribution. A mean value (of any set of observations), is (approximately) normally distributed. If you take, say, 50 random samples out of a certain population each of, say, 25 single observations, the 50 mean values will be (approximately) normally distributed. This result is also valid for the mean values of log-normal distributions, delta distributions or any other type of distribution. This means that the mean values of all distributions observed in fishery biology are approximately normally distributed.

If we divide both sides of Eq. 2.2.1 by n (= sample size) we get:

$$Fc(x)/n = \frac{dL}{s\sqrt{2\pi}} \exp(- (x-\bar{x})^2/(2s^2)), \quad x=1,2,3,... \qquad (2.2.2)$$

the new found values, $Fc(x)/n$, will add up to nearly 1.0. Each value indicates the probability that a randomly drawn fish will belong to the corresponding length interval. That is, they can be interpreted as the probability of a randomly drawn fish to belong to the length interval from $x-dL/2$ to $x+dL/2$.

For the nine length intervals of Table 2.2.1 we find:

j	interval	probability
1	10.5-11.5	0.033
2	11.5-12.5	0.069
3	12.5-13.5	0.116
4	13.5-14.5	0.161
5	14.5-15.5	0.181
6	15.5-16.5	0.166
7	16.5-17.5	0.123
8	17.5-18.5	0.075
9	18.5-19.5	0.037
	Total:	0.961

Thus for example, the odds are 181 to 1000 that a randomly drawn fish will be of a length between 14.5 and 15.5 cm. If we had included all length intervals (and not only the nine for which we had observations) the probabilities would have added up to 1.00.

The normal distribution will be used in length frequency analyses in the following chapters, because the length distribution of a single cohort of fish can be described by a normal distribution. As an introduction we shall study some of its aspects.

$Fc(x)/n$ in Eq. 2.2.2 is the "theoretical relative frequency" (or the probability). The counterpart to it is the "observed relative frequency

$$F(x)/n \qquad (2.2.3)$$

The theoretical and the observed relative frequencies for the sample used are shown in columns B and D of Table 2.2.2. Table 2.2.2 also gives the cumulated relative frequencies, where

$$\frac{1}{n} \sum_{u <= x} Fc(u) \qquad (2.2.4)$$

is the "cumulated theoretical relative frequency" (Column C) and

$$\frac{1}{n} \sum_{u <= x} F(u) \qquad (2.2.5)$$

Table 2.2.2 Cumulated relative frequencies based on Tables 2.1.2 and 2.2.1 (compare Figure 2.2.2)

	A	B	C	D	E
index	length interval x-dL/2 x+dL/2	theoretical relative frequency (Eq.2.2.2)	cumulated theoretical relative frequency (Eq.2.2.4)	observed relative frequency (Eq.2.2.3)	cumulated observed relative frequency (Eq.2.2.5)
tail	0.0-10.5	0.019	0.019	0	0
1	10.5-11.5	0.033	0.052	0.037	0.037
2	11.5-12.5	0.069	0.121	0.111	0.148
3	12.5-13.5	0.116	0.237	0.111	0.259
4	13.5-14.5	0.161	0.398	0.148	0.407
5	14.5-15.5	0.181	0.579	0.148	0.556
6	15.5-16.5	0.166	0.745	0.185	0.741
7	16.5-17.5	0.123	0.868	0.111	0.852
8	17.5-18.5	0.075	0.943	0.074	0.926
9	18.5-19.5	0.137	0.980	0.074	1.000
tail	19.5-∞	0.020	1.000	0.000	1.000

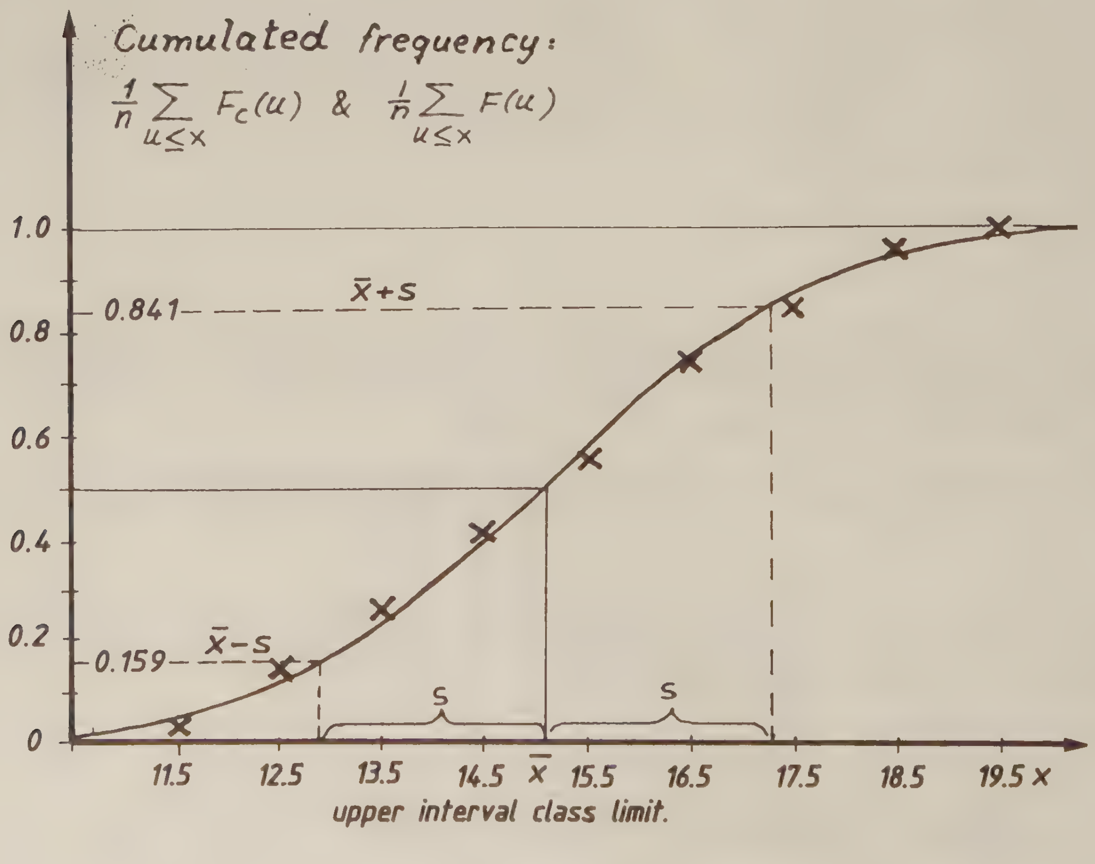


Figure 2.2.2 Theoretical (curve) and observed (crosses) relative cumulated frequencies, based on Table 2.2.2

is the "cumulated observed relative frequency" (Column E). These cumulated frequencies are also shown in Figure 2.2.2. To make the theoretical relative frequencies add up to 1.0 the probabilities of the extremes or "tails" of the normal distribution have also been calculated. Note that in Figure 2.2.1 on the graph for the frequencies (Fc and F on x) the variable x is the interval midpoint (cf. Fig. 2.2.1), whereas in Figure 2.2.2 in the plot of cumulated frequencies on length the variable x is the upper limit of the length interval.

A special property of the normal distribution is that there is the probability 0.159 (or the odds are 159 to 1000) that an observation is less than $\bar{x}$ - s and the same probability that it is greater than $\bar{x}$ + s. This may also be expressed (see Figure 2.2.2) as:

$$\frac{1}{n}\,\Sigma\,Fc(u) = 0.159$$

where $u <= \bar{x}-s$ and

$$\frac{1}{n}\,\Sigma\,Fc(u) = 0.841\ (= 1 - 0.159) \qquad (2.2.6)$$

where $u <= \bar{x}+s$

The standard deviation, s, can be estimated graphically by utilizing this property of the normal distribution as indicated in Figure 2.2.2.

The procedures to calculate the mean and the standard deviation (Table 2.1.2) can be performed on any length frequency data set. However, if for some reason, the observed frequency diagram does not represent the entire distribution, then the obtained values (from Eqs. 2.1.1 and 2.1.2) for sample mean and variance will be biased, i.e. the sample mean and variance may have no relation to the population mean and variance. The concept of "bias" will be further discussed in Section 2.7. If, for example, only the frequencies in the length interval from 10 to 15 cm are available (i.e. only the data for the left hand side) we are in a situation where Eq. 2.1.1 (mean value) and Eq. 2.1.2 (variance) do not represent the population. As will appear in Chapter 3 this is often the case when analyzing length frequencies. However, there are a number of methods to overcome the problem.

A simple graphical method is based on the use of "normal distribution paper", or "probability paper", which is constructed so that a cumulated normal distribution appears as a straight line when plotted. Probability paper may be used to visually check if the observed cumulated frequency conforms to a normal distribution. It is difficult by visual inspection to assess whether points conform to a bell-shaped or a sigmoid curve, whereas it is relatively easy to see if points conform to a straight line. Figure 2.2.3 illustrates the transformation of the cumulated relative frequency curve from Figure 2.2.2 into a straight line using normal distribution graph paper.

Figure 2.2.4 shows the observed and theoretical cumulated frequencies of Table 2.2.2 plotted on probability paper. The line in Figure 2.2.4 is fitted to eight points. Note, however, that since it is a straight line it can also be estimated in a case where only the frequencies of the first, say, five length groups are available. To calculate the first five relative frequencies we must know the first five absolute frequencies and the total number of observations. The mean value is estimated graphically by the intersection between the line that fits the points and the horizontal line drawn through the 50% point. The standard deviation, s, can also be estimated graphically by means of the 16% point (or the 84% point), as demonstrated in Figure 2.2.4, as a concequence of the properties described by Eq. 2.2.6.

(See **Exercise(s)** in Chapter 17).

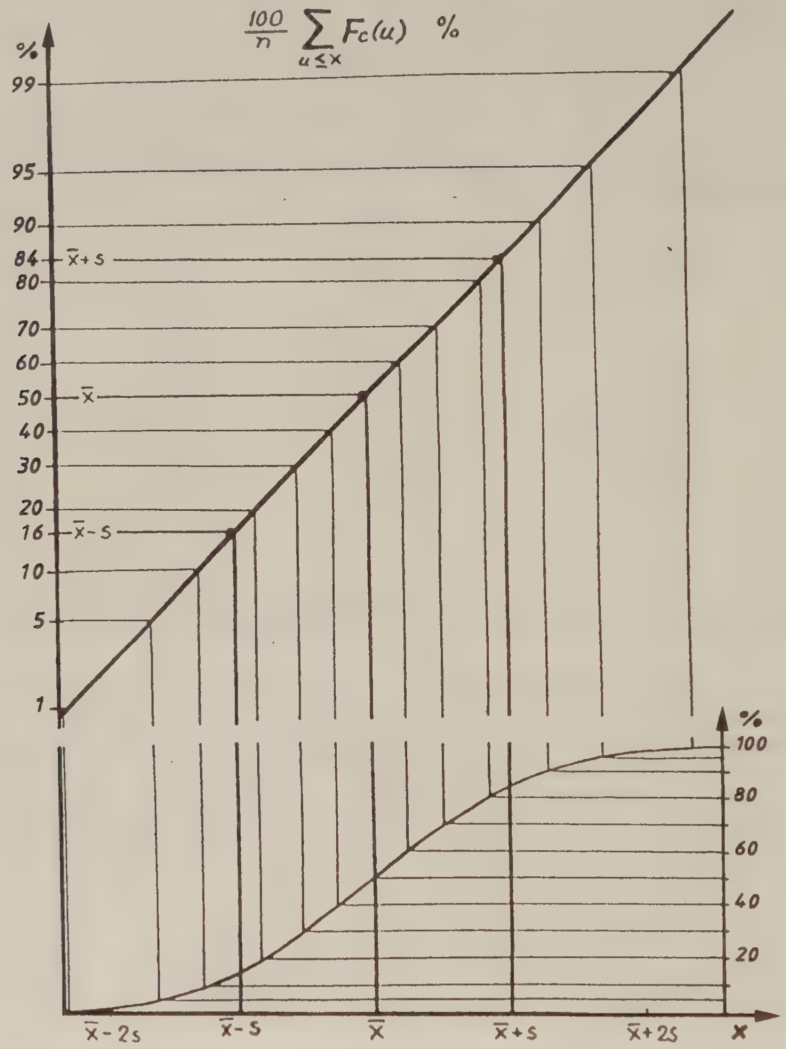


Figure 2.2.3 Transformation of the cumulated normal distribution curve into a straight line

2.3 Confidence limits

In this section we shall also use the example of a length composition sample of fish from one cohort. We have estimated the mean length of the cohort, $\bar{x}$, from the sample. Such an estimate is usually different from the true population mean, the mean we would have obtained if all fish of that cohort in the sea had been measured. Usually the true mean length is unknown. In case we were dealing with a population of cultured fish in a pond we might be able to measure the true mean length of that population, but for a wild fish stock it is impossible to measure the true value of any parameter. In practice this also applies to the population of fish caught in a fishery, since we will not be in a position to measure all fish caught. We shall deal with the precision of the estimate of the mean length, in other words how great the deviation between the estimate and the true mean is likely to be. This uncertainty about the true mean is

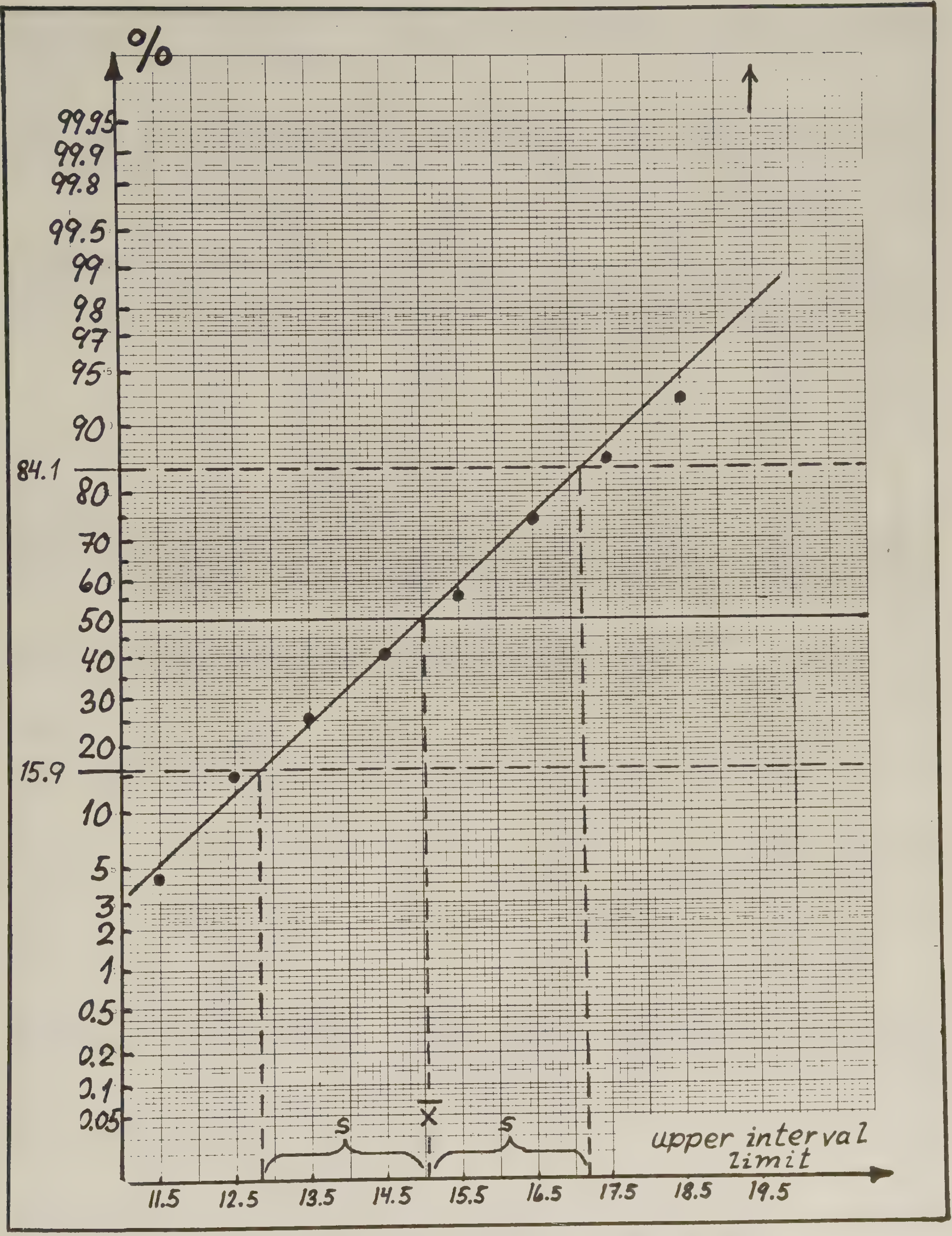


Figure 2.2.4 Observed and theoretical cumulated relative frequencies plotted on probability paper. Note that the points are plotted against the upper interval limits. (Based on data from Table 2.2.2)

expressed by the "confidence limits". In the case of a normal distribution, such confidence limits are given by:

$$L1 = \bar{x} - t(n-1)\ s/\sqrt{n} \quad \text{and} \quad L2 = \bar{x} + t(n-1)\ s/\sqrt{n} \qquad (2.3.1)$$

where n is the sample size and t(n-1) are the so-called fractiles in the "Student's t-distribution" (Table 2.3.1). The argument "f" in the t-distribution (Table 2.3.1) is called the "number of degrees of freedom". In general the number of degrees of freedom is the number of observations minus the number of parameters. Here, $\bar{x}$ is the only parameter.

Returning to the example given in Section 2.1 (Table 2.1.2) we want to calculate, for example, the 95% confidence limits for the mean length of fish in the population from which the sample was drawn. We use the 95% fractile of the t distribution (Table 2.3.1) with n-1 = 26 degrees of freedom and insert into Eq. 2.3.1:

$$t(n-1)\ s/\sqrt{n} = 2.06 * 2.20/\sqrt{27} = 0.87$$

The confidence limits then are:

lower limit: $\bar{x} - 0.87 = 15.07 - 0.87 = 14.20$
upper limit: $\bar{x} + 0.87 = 15.07 + 0.87 = 15.94$

Thus, we are "95% confident" that the true mean length lies somewhere between 14.20 and 15.94, or in other words, if sampling was repeated 100 times under the same conditions we would expect the means to lie 95 times between 14.20 and 15.94. The interval between the lower limit and the upper limit is called the "confidence interval".

The quantity $s/\sqrt{n}$ is the standard deviation of the estimate of the mean length (also called the " standard error") so that $\bar{x}$ has the variance (compare to Table 2.1.1):

$$VAR(\bar{x}) = s^2/n \qquad (2.3.2)$$

Thus, the larger the sample, the more precise is the estimate of $\bar{x}$ (this subject will be discussed further in Section 2.8).

Eq. 2.3.2 follows from two general rules for random variables which are applied repeatedly in this manual. They are:

$$VAR\ (Cx) = C^2 VAR(x) \qquad (2.3.3)$$

$$VAR\ (\sum_{i=1}^{n} x) = n\ VAR(x) \qquad (2.3.4)$$

where C is a constant. For instance, when the variance of x is s^2 then the variance of 3x is $9s^2$; or, when the original observations are summed three by three, then the variance of $x_1+x_2+x_3$ is $3\ s^2$.

The above statements about confidence limits apply only to "unbiased" estimates of the mean value. In cases when samples are biased, no matter how many fish we sample and measure we shall always get estimates of the mean value which are different from the true mean value.

Suppose we want to estimate the mean length of a certain fish species caught in a commercial fishery (note: fish caught are the fish landed plus the fish discarded at sea). Thus, if we sample only from the landings, and not the fish, usually below a certain size, which are discarded at sea, we get a biased estimate of the mean length of the fish caught. The mean length of the catch will be overestimated, no matter how many fish we sample at the landing site. We only can get an unbiased estimate of the mean length of the landed fish.

(See **Exercise(s)** in Chapter 17).

Table 2.3.1 Fractiles of the t-distribution

degrees of freedom	fractiles			degrees of freedom	fractiles		
f *)	90% t(f)	95% t(f)	99% t(f)	f	90% t(f)	95% t(f)	99% t(f)
1	6.31	12.71	63.66	15	1.75	2.13	2.95
2	2.92	4.30	9.93	16	1.75	2.12	2.92
3	2.35	3.18	5.84	17	1.74	2.11	2.90
4	2.13	2.78	4.60	18	1.73	2.10	2.88
5	2.02	2.57	4.03	19	1.73	2.09	2.86
6	1.94	2.45	3.71	20	1.73	2.09	2.85
7	1.90	2.37	3.50	25	1.71	2.06	2.79
8	1.86	2.31	3.36	30	1.70	2.04	2.75
9	1.83	2.26	3.25	40	1.68	2.02	2.70
10	1.81	2.23	3.17	50	1.67	2.01	2.68
11	1.80	2.20	3.11	60	1.67	2.00	2.66
12	1.78	2.18	3.06	80	1.67	1.99	2.64
13	1.77	2.16	3.01	100	1.66	1.98	2.63
14	1.76	2.15	2.98	∞	1.65	1.96	2.58

*) The use of the letter t in this context is universal. In this manual t is also used to represent the age of a fish

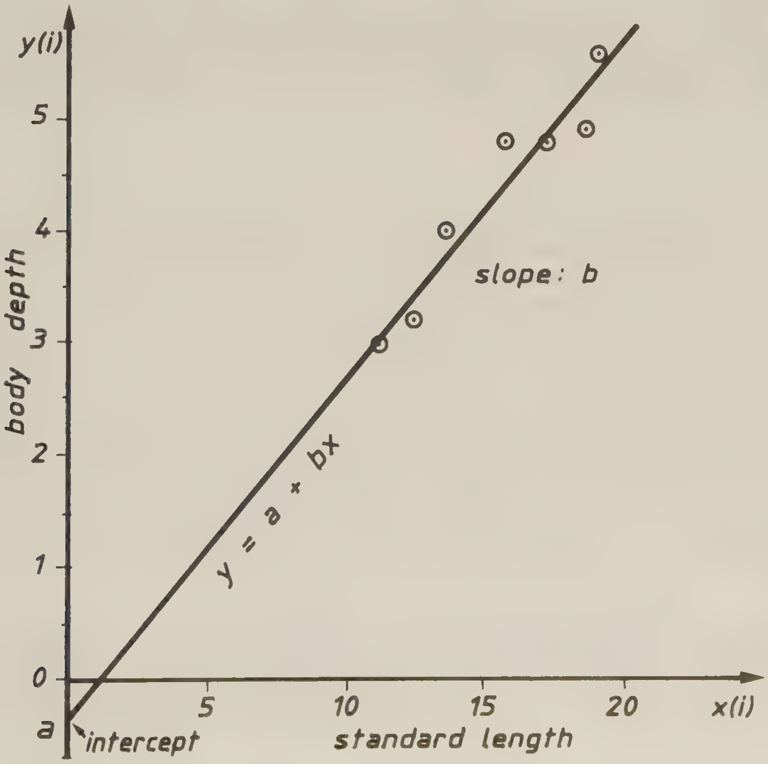


Figure 2.4.1 Scatter diagram of body depth (y) against standard length (x), also called the "plot of y on x"

2.4 Ordinary linear regression analysis

This method is used when we want to describe the variation of one quantity, e.g. the body depth of a fish, as a linear function of another quantity, e.g. the body length. The theory requires that the quantity on the horizontal axis (the independent variable) is measured with absolute precision. The method is often applied, however, when this requirement is violated. The effect of the inaccuracy of the values of the independent variable is that the slope of the line becomes flatter (closer to zero).

Suppose we have measured both the standard length and the body depth of a sample of 7 fish.

Table 2.4.1 shows the standard lengths, x(i), and the body depths, y(i), i = 1,2,...,7.

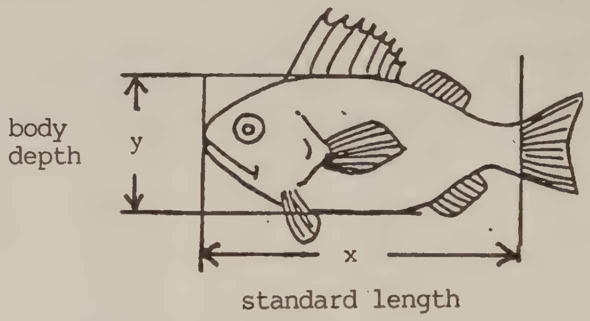


Table 2.4.1 Sample of standard lengths, x, and body depths, y

i	1	2	3	4	5	6	7
x(i)	11.2	12.4	13.5	15.7	17.1	18.5	19.0
y(i)	3.0	3.2	4.0	4.8	4.8	4.9	5.6

As can be expected, the body depth tends to increase when the length increases. If the body proportions of a fish remained constant for all body sizes depth would be proportional to length, and could be described by the model:

$$y(i) = b\ x(i) \qquad (2.4.1)$$

where b is a constant, also called a "parameter". Figure 2.4.1 shows the "plot" (or the "scatter diagram") of y(i) against x(i). We may allow for a deviation from proportionality between x and y by introducing a second parameter, a, and use instead of Eq. 2.4.1 the model:

$$y(i) = a + b\ x(i) \qquad (2.4.2)$$

An implication of Eq. 2.4.2 is that a fish of zero length has depth a, which makes no sense except when a is zero. However, if only lengths in a certain range are considered (e.g. only lengths above 5 cm), the two parameter model may give a better fit to the observations than the one parameter model, because the assumption of proportionality between length and depth is not strictly fulfilled.

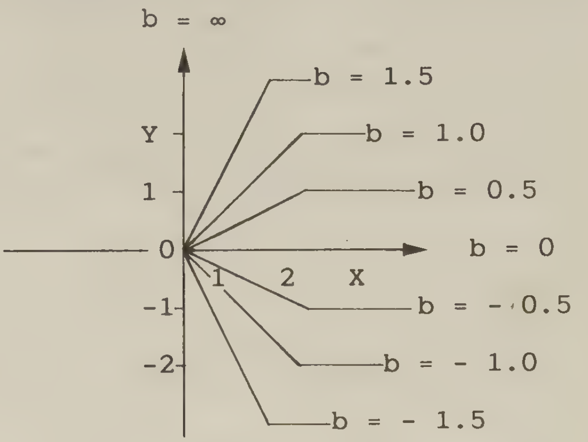
The mathematical model of Eq. 2.4.2 is called a "linear model" because pairs (x,y) which conform to the model, lie on a straight line. With a = -0.32 and b = 0.30 we get the straight line shown in Figure 2.4.1. With these values of a and b the line in Figure 2.4.1 fits well to the observed pairs of (x,y).

We shall now look into the problem of determining the line, i.e. how to estimate the parameters a and b. Just as we did for the mean value (cf. Section 2.3) we shall also show how the confidence limits of the estimates of a and b are calculated. This procedure is called "ordinary linear regression analysis". This method is probably the most commonly used statistical technique in fishery biology. There are special names for the parameters: a is called the "intercept" and b is called the "slope". The intercept is the distance from the point (0.0) in the (x,y)-diagram to the point where the "regression line"

$$y = a + bx$$

intersects with the y-axis (see Figure 2.4.1). Sometimes a is called the "Y-intercept" while the distance from (0,0) to the point where the line intersects the x-axis is called the "X-intercept".

The slope indicates how steep the line is. If b = 0 the line is parallel to the X-axis. If b is positive the slope is ascending. If b is negative the slope is descending.



The variable on the horizontal axis, x, is called the "independent variable" and the variable on the vertical axis, y, is called the "dependent variable". The regression line is determined as the line which minimizes the sum of squares of deviations between the line y = a + b x and the pairs of observations, (x(i),y(i)). We say that a and b are estimated by the "least squares method", i.e. we search for the values of a and b which minimize:

$$\sum_{i=1}^{N} (y(i) - a - b\,x(i))^2 \tag{2.4.3}$$

where n is the number of pairs of observations (n = 7 in the example). The deviations between the line and the observations are illustrated in Figure 2.4.2. The assumption behind the regression analysis is that each y(i) is normally distributed with mean value a + b x(i), and with a constant variance, i.e. a variance which is not dependent on the value of x(i). The formula for the estimate of this common variance differs only slightly from the one introduced in Section 2.1. The estimate is:

$$s^2 = \frac{1}{n-2} \sum_{i=1}^{n} (y(i) - a - b\,x(i))^2 \tag{2.4.4}$$

This s^2 is called the "variance about the regression line". There are n-2 degrees of freedom (the number by which the sum is divided) because we have two parameters, a and b.

Table 2.4.2 The calculation procedure for ordinary linear regression analysis. Results marked by *) are not used in the calculation of a and b, but are derived here for subsequent use

	standard length		body depth		
i	$x(i)$	$x(i)^2$	$y(i)$	$y(i)^2$	$x(i)y(i)$
1	11.2	125.44	3.0	9.00	33.60
2	12.4	153.76	3.2	10.24	39.68
3	13.5	182.25	4.0	16.00	54.00
4	15.7	246.49	4.8	23.04	75.36
5	17.1	292.41	4.8	23.04	82.08
6	18.5	342.25	4.9	24.01	90.65
7=n	19.0	361.00	5.6	31.36	106.40
Σ	107.4	1703.6	30.3	136.69	481.77
	$\Sigma x(i)$	$\Sigma x(i)^2$	$\Sigma y(i)$	$\Sigma y(i)^2$	$\Sigma x(i)y(i)$

$\bar{x} = 15.343$	$\bar{y} = 4.329$
$\frac{1}{n} (\Sigma x(i))^2 = 1647.82$	$\frac{1}{n} (\Sigma y(i))^2 = 131.16$ *)
$\Sigma x(i)^2 - \frac{1}{n} (\Sigma x(i))^2 = 55.78$	$\Sigma y(i)^2 - \frac{1}{n} (\Sigma y(i))^2 = 5.534$ *)
$sx^2 = 9.297$ *) $sx = 3.049$ *)	$sy^2 = 0.922$ *) $sy = 0.960$ *)

$\frac{1}{n} \Sigma y(i) \Sigma x(i) = 464.89$

$\Sigma y(i)x(i) - \frac{1}{n} \Sigma y(i)\, \Sigma x(i) = 16.88$ $\quad sxy = 2.814$ *)

$$b = \frac{\Sigma y(i)x(i) - \frac{1}{n} \Sigma y(i)\, \Sigma x(i)}{\Sigma x(i)^2 - \frac{1}{n} (\Sigma x(i))^2} = \frac{16.88}{55.78} = 0.303$$

$a = \bar{y} - \bar{x}\, b = 4.329 - 15.343 * 0.303 = -0.315$

Table 2.4.3 Calculation of variance about the line from Eq. 2.4.4

i	$x(i)$	$y(i)$	$a + b\, x(i)$	$(y(i) - a - b\, x(i))^2$
1	11.2	3.0	3.079	0.0062
2	12.4	3.2	3.442	0.0587
3	13.5	4.0	3.776	0.0504
4	15.7	4.8	4.442	0.1281
5	17.1	4.8	4.866	0.0044
6	18.5	4.9	5.291	0.1525
7	19.0	5.6	5.442	0.0250
			sum:	0.4252

$s^2 = 0.4252/(7-2) = 0.085$

Estimates of the parameters a (intercept) and b (slope) are obtained by:

$$b = \frac{\sum_{i=1}^{n} y(i)x(i) - \frac{1}{n} \sum_{i=1}^{n} y(i) \sum_{i=1}^{n} x(i)}{\sum_{i=1}^{n} x(i)^2 - \frac{1}{n} (\sum_{i=1}^{n} x(i))^2} \qquad (2.4.5)$$

$$a = \bar{y} - \bar{x}\, b \qquad (2.4.6)$$

where $\bar{y}$ and $\bar{x}$ are the mean values of y and x as defined by Eq. 2.1.1.

In Table 2.4.2 the calculation procedures to estimate a and b are demonstrated using the data from Table 2.4.1. Thus, the estimated regression line becomes:

$$y = -\,0.315 + 0.303\, x \qquad (2.4.7)$$

To calculate the confidence limits of a and b we need the sum of squares of deviations of x and y. The variances of x and y are defined by Eq. 2.1.3 as follows:

$$sx^2 = \frac{1}{n-1} (\Sigma x(i)^2 - \frac{1}{n} (\Sigma x(i))^2) \qquad (2.4.8)$$

and a similar expression for sy^2. For use in the next section we introduce:

$$sxy = \frac{1}{n-1} \left[\Sigma y(i)x(i) - \frac{1}{n} \Sigma y(i) \Sigma x(i) \right] \qquad (2.4.9)$$

The procedure for the calculation of variance about the regression line leading to Eq. 2.4.4 is demonstrated in Table 2.4.3. However, the variance about the line can be obtained more easily from sy and sx:

$$s^2 = \frac{n-1}{n-2} (\ sy^2 - b^2\, sx^2) \qquad (2.4.10)$$

Given the results from Table 2.4.2, Eq. 2.4.10 becomes:

$$s^2 = \frac{6}{5} (0.922 - 0.303^2 * 9.297) = 0.085$$

The variances of the estimates of b and a are:

$$sb^2 = \frac{1}{n-2} (\ (sy/sx)^2 - b^2\ ) \qquad (2.4.11)$$

and

$$sa^2 = sb^2 (\frac{n-1}{n}\, sx^2 + \bar{x}^2) \qquad (2.4.12)$$

Given the results from Table 2.4.2 we get:

$$sb^2 = \frac{1}{(7-2)} (\ \frac{0.922}{9.297} - 0.303^2\ ) = 0.00147, \qquad sb = 0.038$$

$$sa^2 = 0.00147\ (\ \frac{7-1}{7}\ 9.297 + 15.343^2) = 0.358, \qquad sa = 0.60$$

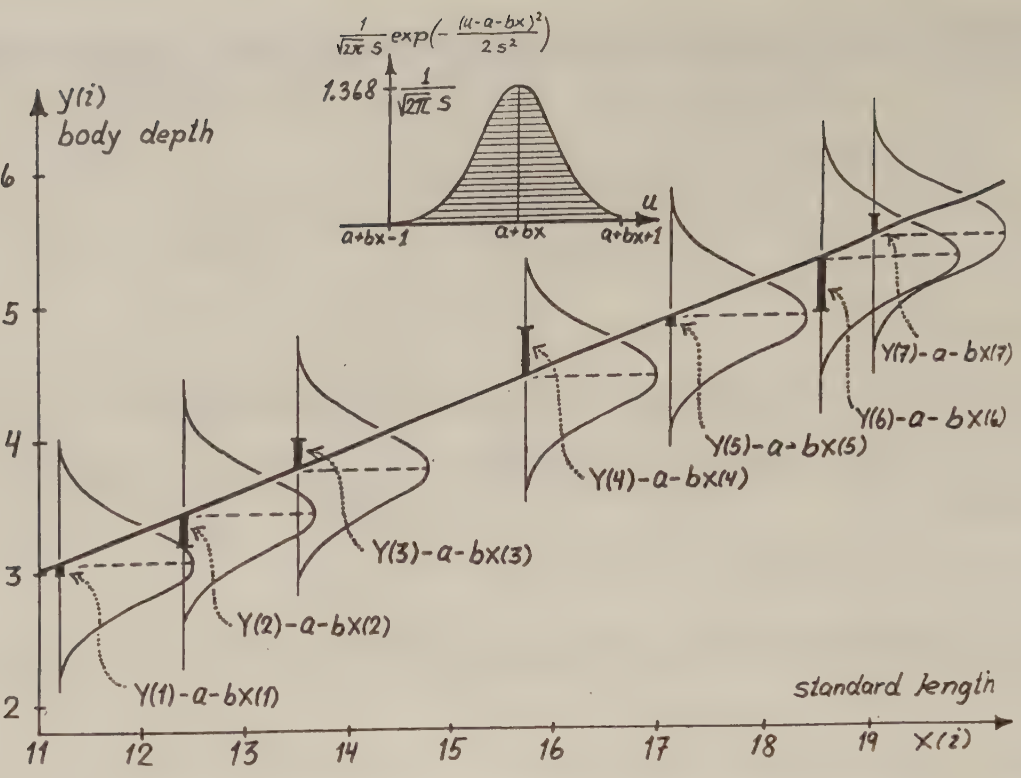

$\frac{1}{\sqrt{2\pi}s}exp(-\frac{(u-a-bx)^2}{2s^2})$
1.368
$\frac{1}{\sqrt{2\pi}s}$
u
a+bx-1
a+bx
a+bx+1
y(i)
body depth
6
5
4
3
2
Y(1)-a-bX(1)
Y(2)-a-bX(2)
Y(3)-a-bX(3)
Y(4)-a-bX(4)
Y(5)-a-bX(5)
Y(6)-a-bX(6)
Y(7)-a-bX(7)
standard length
11
12
13
14
15
16
17
18
19
X(i)

Figure 2.4.2 Illustration of the assumptions behind ordinary linear regression analysis. Each y(i) for a given x(i) is normally distributed with a common variance

The confidence limits for a and b are:

a: [a - sa t(n-2) , a + sa t(n-2)]

b: [b - sb t(n-2) , b + sb t(n-2)]

The 95% confidence limits for the example with n = 7 fish and t(n-2) = 2.57 (Table 2.3.1) become:

a: [-.315 - .60 * 2.57 , -.315 + .60 * 2.57] = [-1.86 , 1.23]

b: [.303 - .038 * 2.57 , .303 + .38 * 2.57] = [0.20 , 0.40]

Note that the confidence interval for the intercept a contains zero. This means that the hypothesis that body depth is directly proportional to length, (thus that "a = 0") cannot be rejected by the 95% confidence limits. We say that a is not significantly different from 0 at the 95% level.

If we have a good reason to assume that a = 0 then the estimated value should be replaced by 0 if the estimate is not significantly different from 0. Then, however, b must be recalculated as follows:

$$b = \Sigma\, x(i)*y(i)/\Sigma x(i)^2 \qquad (2.4.13)$$

Our present estimate is based on only seven fish. If we had measured 200 fish the estimate of the standard deviation, sa, would be smaller (cf. Eqs. 2.4.11 and 2.4.12). Let us assume for example, that $\bar{x}$, $\bar{y}$, sy, a and b were the same for a sample size of n = 200 as those estimated for a sample size of n = 7 (which might well happen). Although the estimates of a and b turn out to have the same value, their standard deviations, sa and sb, will be different.

With n = 200 Eq. 2.4.11 gives sb = 0.006098, while Eq. 2.4.12 gives sa = 0.0091 and t(198) = 1.97 (Table 2.3.1). Thus, sa and sb become smaller, and consequently the confidence interval of a becomes smaller:

$$[-.315-1.97 * 0.0091 \quad , \quad -.315+1.97 * 0.0091] = [-0.33,-0.30]$$

The estimate of a would now be significantly different from 0. In that case we can conclude that the odds are less than 5% that the true value of a is larger than -0.30 or smaller than -0.33.

(See **Exercise(s)** in Chapter 17).

2.5 The correlation coefficient and functional regression

The "correlation coefficient", r, is a measure of the linear association between two quantities, both of which are subject to random variation. The standard length and body depth sample from Section 2.4 is an example of such two quantities. In this case seven fish were drawn at random. By accident we could have drawn seven fish all of (nearly) the same length. In that case the sample would not be suitable for estimation of the length/depth relationship because the confidence limits of a and b would become very wide.

The correlation coefficient can be used only when both measurements are allowed to vary randomly. If we had selected seven fish with predetermined lengths rather than random lengths (e.g. had selected the lengths 4, 6, 8, 10, 12, 14 and 16 cm for the length/depth sample) the calculation of a correlation coefficient for this sample would be incorrect.

The correlation coefficient is defined as:

$$r = \frac{sxy}{sx * sy} \qquad (2.5.1)$$

where sxy is defined by Eq. 2.4.9 and sx and sy are defined by Eq. 2.4.8. Inserting the slope ($b = sxy/sx^2$), Eq. 2.5.1 becomes:

$$r = b\ (sx/sy) \qquad (2.5.2)$$

The range of r is: $-1.0 \le r \le 1.0$. r is negative if y tends to decrease when x increases and r is positive if y tends to increase when x increases. This statement also holds for the slope b and it follows from Eq. 2.5.2: As sx/sy is always positive (cf. the definition Eq. 2.4.8) r has the same sign as the slope b. The extreme cases, r=1 or r=-1 occur when all pairs (x,y) lie exactly on a straight line. The closer r approaches zero the less pronounced is the linear association between y and x. When r = 0, x and y are independent of each other.

Figure 2.5.1 shows four examples of scatter diagrams with different r-values. For the example of Table 2.4.2 we get:

$$r = \frac{2.814}{3.049 * 0.960} = 0.961$$

Let us call r1 and r2 the 95% confidence limits for r. They can be calculated from the expression:

$$\left.\begin{matrix} r2 \\ \\ r1 \end{matrix}\right\} = \tanh\left(0.5 \ln\left(\frac{1+r}{1-r} \right) \pm 1.96/\sqrt{n-3} \right) \qquad (2.5.3)$$

"tanh" is the "hyperbolic tangent" which is standard on many scientific pocket calculators.

With r from the example (r = 0.961, n=7) the 95% confidence limits become: [r1,r2] = [0.75 , 0.99]. The 99% confidence limits can be obtained by replacing the number 1.96 by 2.58 in Eq. 2.5.3.

Often we are interested in knowing whether zero lies in the confidence interval, viz. what the odds are that the linear association is due to chance. In this example the odds are less than 5% that the linear association is due to chance, because zero is not in the confidence interval.

In the example of regressing body depth on standard length, the length was chosen as the independent variable and the body depth as the dependent variable. However, there is no special reason for this choice. Our sample consists of seven randomly chosen fish. We did not control what their lengths and body depths would be, thus, we could as well have made the opposite choice for dependent and independent variable.

One of the assumptions behind linear regression analysis is that the independent variable cannot be a random variable. The independent variable must be something of which we can determine the values beforehand. For example, if the independent variable is the time the sample is taken, it can be determined beforehand. We could decide to collect a sample on the first day of each month. If we measure time in units of years and start with time zero on the first of January, the independent variable would take the values: 0, 1/12, 2/12, 3/12 ... etc. These values are clearly not random variables.

In the case of the seven fish chosen at random in the example above the situation is that because they were chosen at random out of a normal distribution of fish lengths a correlation analysis can be performed on them. On the other hand, we are able to decide on the lengths beforehand. We could choose the four smallest fish and the three biggest. We could also decide, as we did, to take them as they come. Only in the latter case it is permissible to do both kinds of analysis. In the former case only regression analysis will do. On the other hand, it probably would be a more effective way of doing the regression analysis because of the great distance between the observations on the horizontal axis. This would cause a small variance of the slope. Choosing the fish at random, most of them are likely to be medium-sized and contribute little to the determination of the slope which might show a large variance.

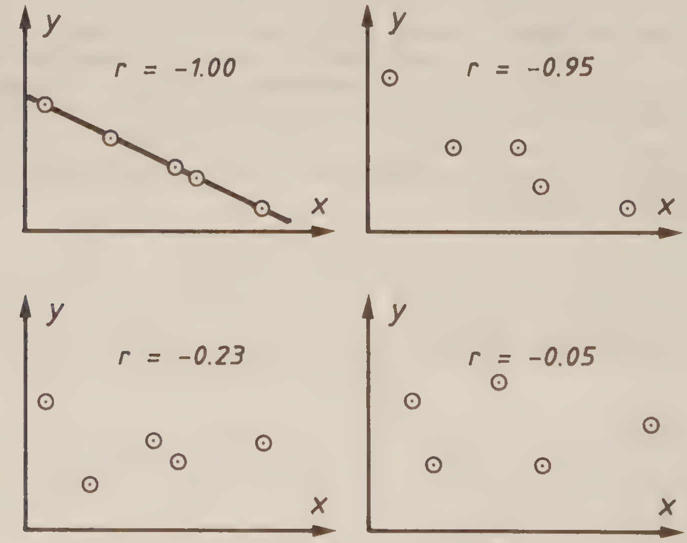


Figure 2.5.1 Examples of correlation coefficients

Another question is whether we should have obtained a different result using the body depth as the independent variable, thus plotting fish length as a function of body depth. First it must be considered whether depth is as precisely measured as length. If it is not, the slope would be biased (flattened) as already mentioned. However, there are problems even if the two variables are measured with the same accuracy.

Taking now the body depth as the independent variable we get what is called an "inverse regression". Only in the exceptional case that all observations lie on the regression line (i.e. if $r = 1$ or $r = -1$) the same result would be obtained for the inverse regression as for the ordinary regression. The equation $y = a + b*x$ (Eq. 2.4.2) is mathematically equivalent to:

$$x = -a/b + y/b$$

or $x = A + B * y$ where $A = -a/b$ and $B = 1/b$ (2.5.4)

Carrying out the inverse regression (Eq. 2.5.4) we find that

$$A = 2.139 \quad \text{and} \quad B = 3.05$$

The equation: $x = 2.139 + 3.05 * y$ can be converted into:

$$y = -0.701 + 0.328\,x$$

which can be compared with the result found for the original regression (Eq. 2.4.7: $y = -0.315 + 0.303 * x$). Thus, the inverse regression gives a result that differs from that of the original regression analysis.

One way to circumvent the problem of choosing the independent variable when both variables are random variables is to use the so-called "functional regression analysis" (see Ricker, 1973). This method estimates a slope (which we call b' to distinguish it from the slope b of the ordinary regression analysis) by the expressions:

$$b' = \begin{cases} sy/sx & \text{if } r > 0 \\ -\,sx/sy & \text{if } r < 0 \end{cases} \qquad (2.5.5)$$

and the intercept:

$$a' = \bar{y} - b' * \bar{x} \qquad (2.5.6)$$

This type of analysis gives a result that may be considered a compromise between the original ordinary regression and its inverse counterpart.

With the results from Table 2.4.2 we get:

$$b' = 0.960/3.049 = 0.315 \text{ and } a' = 4.329 - 0.315 * 15.343 = -0.504$$

and $y = -0.504 + 0.315x$

Functional regression analysis is mentioned here for the sake of completeness. There are some rather intricate limitations to its applicability which we cannot go into here.

The following three regression lines have now been estimated:

1. Original ordinary regression analysis: $y = -0.315 + 0.303 * x$
2. Functional regression analysis : $y = -0.504 + 0.315 * x$
3. Inverse ordinary regression analysis : $y = -0.701 + 0.328 * x$

Figure 2.5.2 shows the three regression lines. Note that all three lines pass through the point $(\bar{x},\bar{y})$ and that an increase in slope is partly balanced by a decrease of the intercept.

(See **Exercise(s)** in Chapter 17).

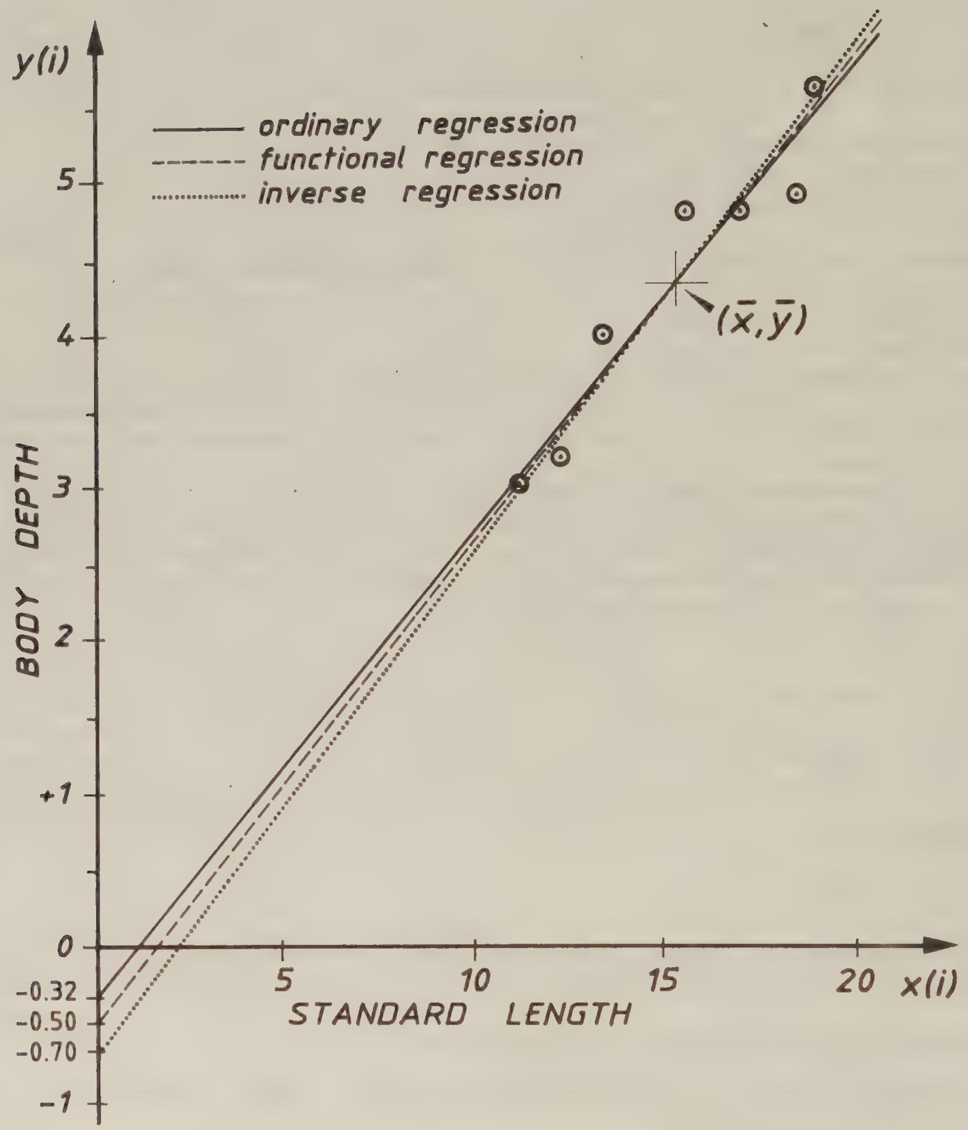


Figure 2.5.2 Functional and inverse regression lines compared to the original regression line

2.6 Linearizing transformations

Linear functions are mathematically easy and also have the advantage that they can be graphically interpreted without any problem. However, many functional relationships observed in fishery biology are not linear. Fortunately, such non-linear functions can often be transformed into linear functions, which means that they can be dealt with in the way described in the foregoing sections. Several examples are given below of the application of transformations from non-linear functions into linear functions in fishery biology.

EXAMPLE 1: LENGTH-WEIGHT RELATIONSHIP

Here we consider a famous example, namely the functional relationship between body length and body weight of fish. Figure 2.6.1 shows a plot of weight on length of the threadfin bream, Nemipterus marginatus. Clearly, this is not a linear relationship. The curve in Figure 2.6.1 is of the function:

$$W(i) = q\ L(i)^b \tag{2.6.1}$$

where W(i) is the body weight of fish no. i, L(i) is the body length and q and b are parameters. Eq. 2.6.1 is usually called the "length-weight relationship" and can be transformed into a linear equation by taking logarithms on both sides:

$$\ln W(i) = \ln q + b \ln L(i) \tag{2.6.2}$$

or

$$y(i) = a + b\ x(i) \tag{2.6.2a}$$

where $y(i) = \ln W(i)$, $x(i) = \ln L(i)$ and $a = \ln q$.

With Eq. 2.6.2a we are now in a position to carry out the estimation of a and b by linear regression analysis. Input data are shown in Table 2.6.1 and the corresponding scatter diagram in Fig. 2.6.2. The results are:

$$a\ (= \ln q) = -4.538, \quad b = 3.057 \quad \text{and} \quad r = .996$$

Using Eq. 2.4.11 we can also calculate the confidence limits of b:

$$sb^2 = \frac{1}{16-2}\left(\left(\frac{1.0161}{0.3311}\right)^2 - 3.057^2\right) = 0.0052 \qquad sb = 0.072$$

$sb * t(n-2) = 0.072 * 2.15 = 0.155$ so that the 95% confidence interval for b becomes: [2.90 , 3.21]. The confidence limits tell us that only the first decimal in the estimate of b is significant (cf. Section 2.3), thus the true value of b could as well be 3.0. As the weight of a fish (in grammes) is approximately equal to its volume (in cubic cm) and as its volume is often proportional to L^3 we would expect a value of b close to 3.0 in Eq. 2.6.1 and Eq. 2.6.2. From the confidence limits of b we cannot reject the hypothesis that b = 3.0, and since the argument above also supports this hypothesis we replace the estimate b = 3.057 by 3.0, and obtain a new estimate of the intercept:

$$a = \ln q = \bar{y} - 3\ \bar{x} = 3.799 - 3 * 2.727 = -4.382$$

Table 2.6.1 Data for estimation of a length-weight relationship for the threadfin bream (Nemipterus marginatus) from the South China Sea (from Pauly, 1983)

i	L(i)	W(i)	ln L(i) x(i)	ln W(i) y(i)
1	8.1	6.3	2.092	1.841
2	9.1	9.6	2.208	2.262
3	10.2	11.6	2.322	2.451
4	11.9	18.5	2.477	2.918
5	12.2	26.2	2.501	3.266
6	13.8	36.1	2.625	3.586
7	14.8	40.1	2.695	3.691
8	15.7	47.3	2.754	3.857
9	16.6	65.6	2.809	4.184
10	17.7	69.4	2.874	4.240
11	18.7	76.4	2.929	4.336
12	19.0	82.5	2.944	4.413
13	20.6	106.6	3.025	4.669
14	21.9	119.8	3.086	4.786
15	22.9	169.2	3.131	5.131
16	23.5	173.3	3.157	5.155
		sum	43.629	60.786
		mean	2.7268	3.7991
		sx and sy	0.3311	1.0161

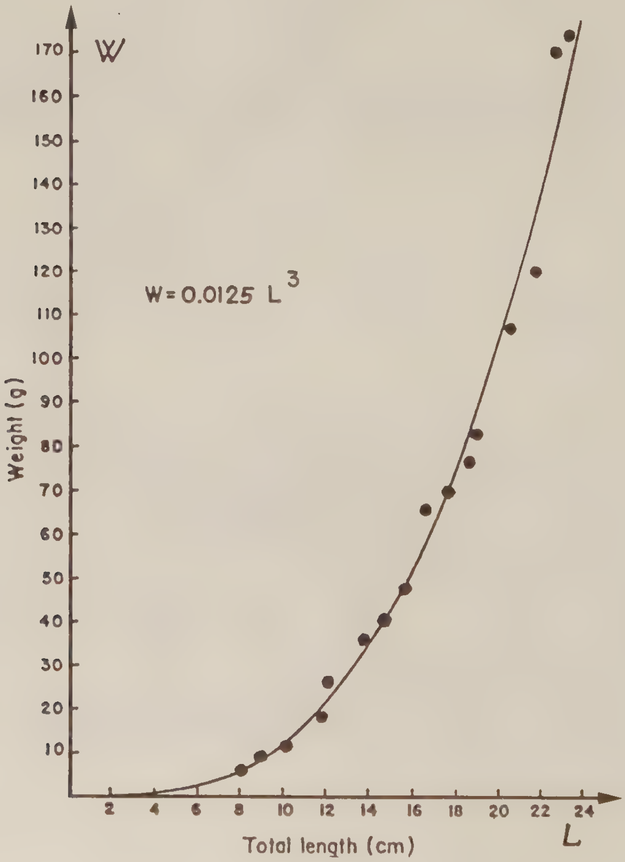


Figure 2.6.1 Length-weight relationship of Nemipterus marginatus in the South China Sea. (Based on data from Table 2.6.1)

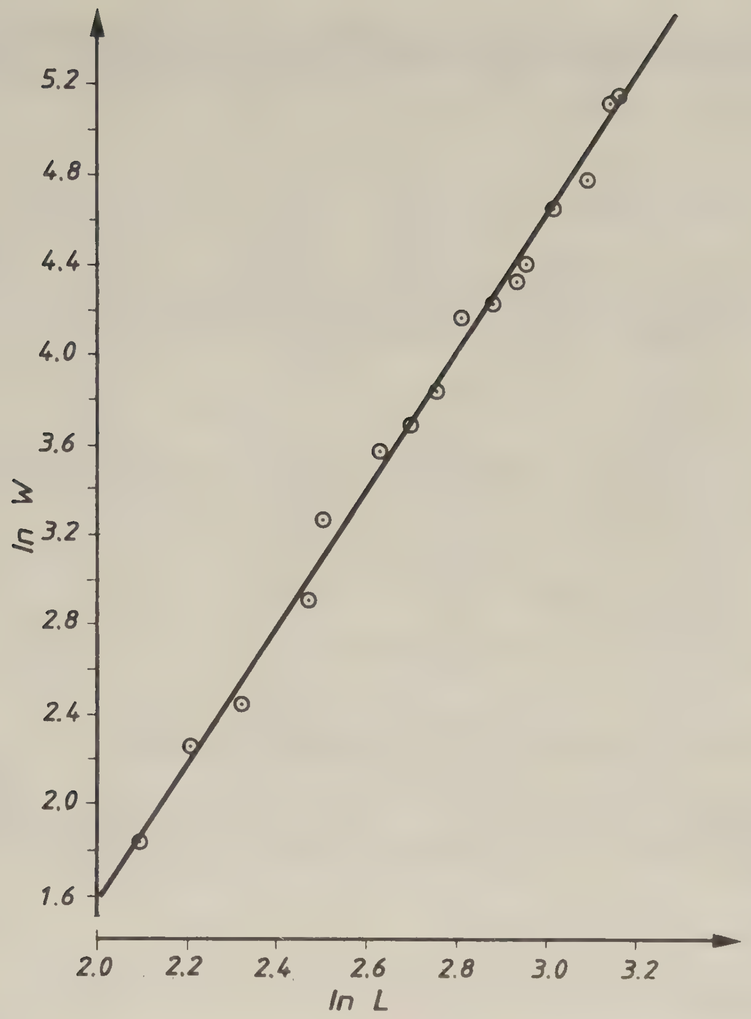


Figure 2.6.2 The data from Figure 2.6.1 converted to natural logarithms

We can now return to the original relationship Eq. 2.6.1 by calculating q from a (=ln q).

$$q = \exp(-4.382) = 0.0125$$

Thus the estimated relationship between W (in grammes) and L (in centimeters) becomes:

$$W = 0.0125 \; L^3$$

The back-transformation from logarithms introduces a bias which we shall not go into here.

EXAMPLE 2: LINEARIZATION OF A NORMAL DISTRIBUTION

In Section 2.2 (Eq. 2.2.1) the mathematical expression for a normal distribution is given as:

$$Fc(x) = \frac{n \; dL}{s \sqrt{2\pi}} \exp \left[-(x - \bar{x})^2/(2 \; s^2) \right]$$

This equation can be transformed into a linear regression in the following two steps:

Step 1: Converting a normal distribution into a parabola

Taking the logarithms on both sides of Eq. 2.2.1 gives:

$$\ln Fc(x) = \ln \frac{n\ dL}{s\ \sqrt{2\pi}} - (x - \bar{x})^2 / (2\ s^2) \qquad (2.6.3)$$

Considering ln Fc(x) as the dependent variable, y, and x as the independent variable, we have hereby obtained a functional relationship between y and x, which can graphically be represented by a parabola which has the general formula:

$$y = a + bx + cx^2$$

Inserting the values used in the example of Table 2.1.2 we obtain:

$$y = \ln (9 * 1/ 2.2\sqrt{2\pi}) - (x - 15.07)^2/(2 * 4.84) =$$

$$1.59 - (x - 15.07)^2/9.68$$

the graph of which is shown in Figure 2.6.3.

Step 2: Converting a parabola into a straight line

In order to transform the parabola into a straight line we have to introduce a new dependent variable, y', which is the difference between the logarithms of the number in a certain length class and the number in the preceding class.

$$y' = \ln Fc(x+dL) - \ln Fc(x) \qquad (2.6.4)$$

This can also be expressed as

$$y' = \Delta \ln Fc(x+dL/2)$$

where Δ (delta) designates a "small" difference between two function

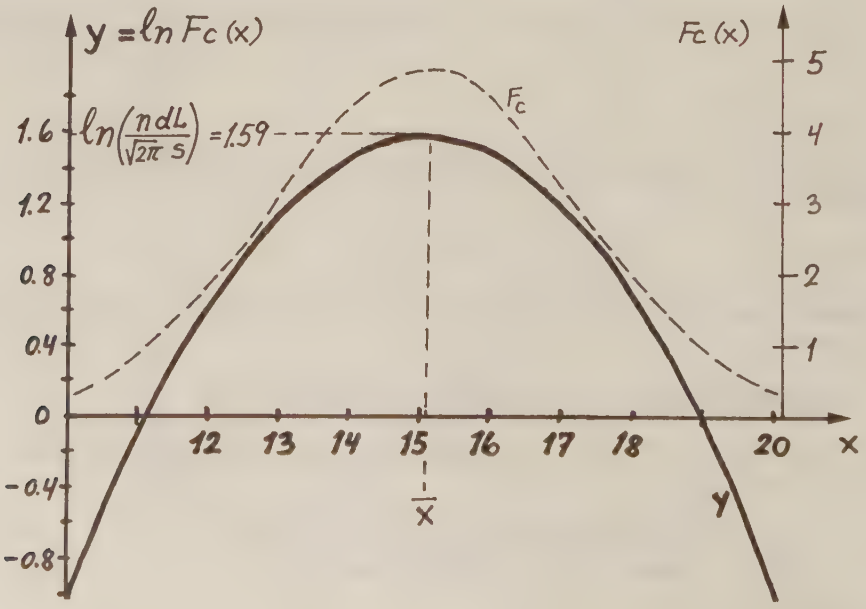


Figure 2.6.3 The ln-transformed normal distribution (y) together with the original distribution (Fc)

Table 2.6.2 Estimation of mean value and variance from the Bhattacharya plot, illustrated by the theoretical frequencies of Table 2.1.1

index	x	x-dL/2,x+dL/2	Fc(x)	y ln Fc(x)	y' /\ln Fc(z)	z x+dL/2
1	11	10.5-11.5	0.88	-0.128	0.743	11.5
2	12	11.5-12.5	1.85	0.615	0.529	12.5
3	13	12.5-13.5	3.14	1.144	0.326	13.5
4	14	13.5-14.5	4.35	1.470	0.117	14.5
5	15	14.5-15.5	4.89	1.587	-0.088	15.5
6	16	15.5-16.5	4.48	1.500	-0.297	16.5
7	17	16.5-17.5	3.33	1.203	-0.500	17.5
8	18	17.5-18.5	2.02	0.703	-0.713	18.5
9	19	18.5-19.5	0.99	0.010		

$b = -0.2073$ $\quad a = 3.1237$

$\bar{x} = -a/b = 15.07$ $\quad s^2 = -dL/b = 4.82$ $\quad (dL = 1)$

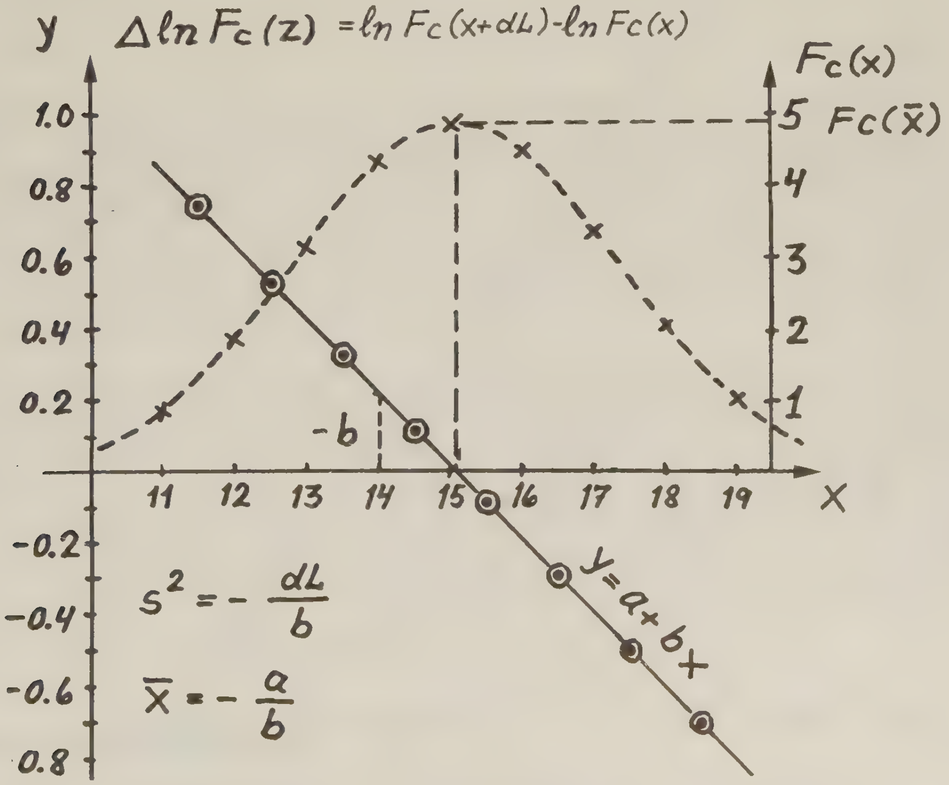


Figure 2.6.4 Bhattacharya plot together with the original normal distribution. (Based on Table 2.6.2)

Table 2.6.3 Bhattacharya plot corresponding to Table 2.1.2

	x			y	y'	z
index	x	x-dL/2,x+dL/2	F(x)	ln Fc(x)	∆ln Fc(z)	x+dL/2
1-2	11.5	10.5-12.5	4	1.386	0.560	12.5
3-4	13.5	12.5-14.5	7	1.946	0.251	14.5
5-6	15.5	14.5-16.5	9	2.197	-0.588	16.5
7-8	17.5	16.5-18.5	5	1.609	-0.916	18.5
9	19.5	18.5-20.5	2	0.693		

$b = -0.263$ $a = 3.909$

$\bar{x} = -a/b = 14.8$ $s^2 = -dL/b = 7.6$ $(dL = 2)$

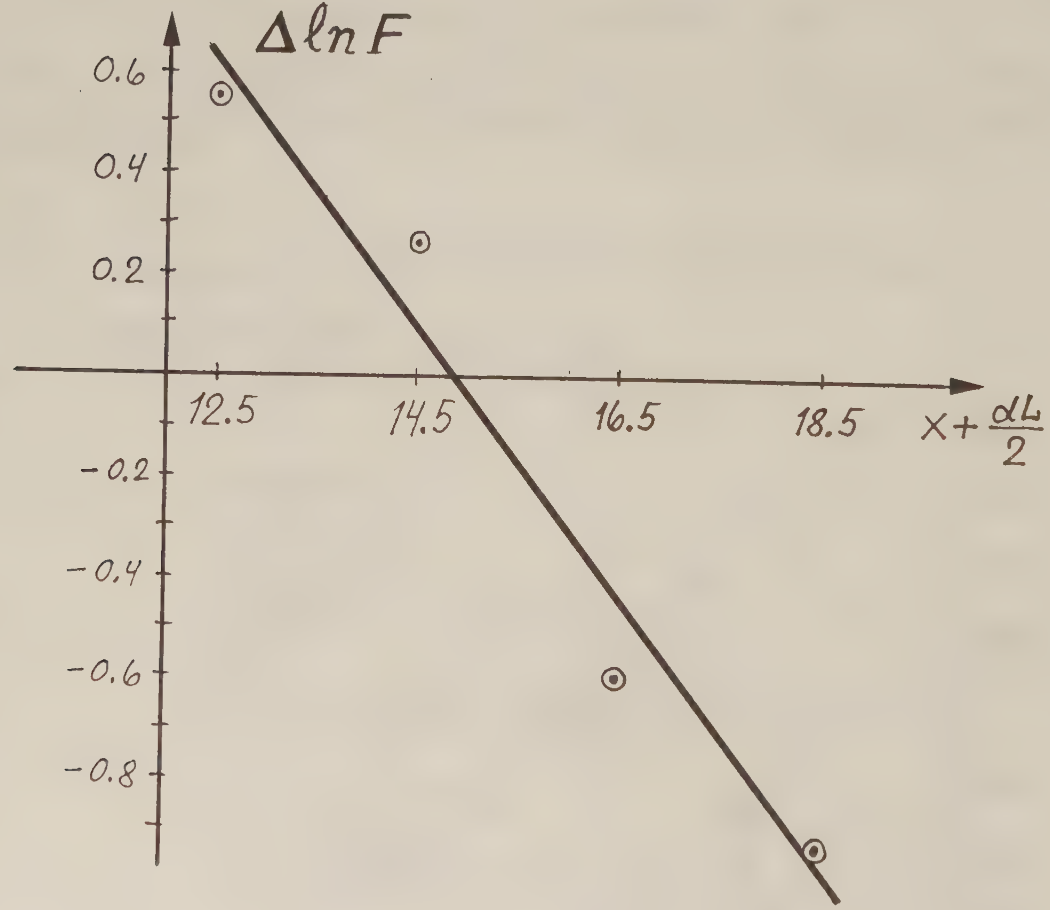


Figure 2.6.5 Bhattacharya plot corresponding to Table 2.6.3

values. y' is to be plotted against a new independent variable, z, which is equivalent to x plus half the length interval:

$$z = x + dL/2$$

We now have to insert Eq. 2.6.3 into Eq. 2.6.4 as follows:

$$y' = \Delta \ln Fc(x+dL/2) = \Delta \ln Fc(z) =$$

$$= \left[\ln \frac{n\ dL}{s\ \sqrt{2\pi}} - \frac{(x+dL-\bar{x})^2}{2\ s^2} \right] - \left[\ln \frac{n\ dL}{s\ \sqrt{2\pi}} - \frac{(x-\bar{x})^2}{2\ s^2} \right]$$

$$= (1/(2s^2)) \left[-(x+dL-\bar{x})^2 + (x-\bar{x})^2 \right]$$

After squaring and summing this can be converted into the relatively simple equation in which for simplicity we again call the dependent variable y:

$$y = \frac{dL\ \bar{x}}{s^2} - \frac{dL\ (x + dL/2}{s^2} \qquad (2.6.5)$$

or $y = a + b\ z$ with $a = dL\ \bar{x}/s^2$, $b = -\ dL/s^2$ and $z = x + dL/2$.

From the slope, b, and the intercept, a, we get the variance:

$$s^2 = -\ dL/b \qquad (2.6.6)$$

and the mean value by

$$\bar{x} = -\ a/b \qquad (2.6.7)$$

This regression is one of the main elements of the method described by Bhattacharya (1967) for separating two or more normal distributions (Section 3.4.1). We call it the "Bhattacharya plot". Table 2.6.2 and Fig. 2.6.4 show an example. In this case the theoretical values, Fc, from Table 2.2.1 have been used as "observations". These conform exactly to the model and mean and variance estimated by the Bhattacharya plot should be the same as those obtained by the traditional method (as in Table 2.1.2). Figure 2.6.4 shows the plot of differences of logarithms against the upper limit of the smallest length group.

Table 2.6.3 shows the estimation of mean and variance by the Bhattacharya plot, but now with the actual observations given in Table 2.1.2. Because of the small sample size the observations have been grouped into two centimeter intervals. Figure 2.6.5 shows the corresponding plot. The estimates of mean and variance obtained in Table 2.6.3 deviate from those calculated by the traditional method (Table 2.1.2) because of the small sample size, because of the bias introduced by large length intervals and because a different statistical method (linear regression analysis) was used.

(See **Exercise(s)** in Chapter 17).

2.7 Simple random sampling

Let us go back to the problem of estimating the mean length of a cohort as addressed already in Sections 2.2 and 2.3. An estimate is said to be "unbiased" if replicate estimates deviate from the true value in a random manner only. The "true value" is the parameter value we would get from measuring all specimens in the total population (cf. Section 2.3). An

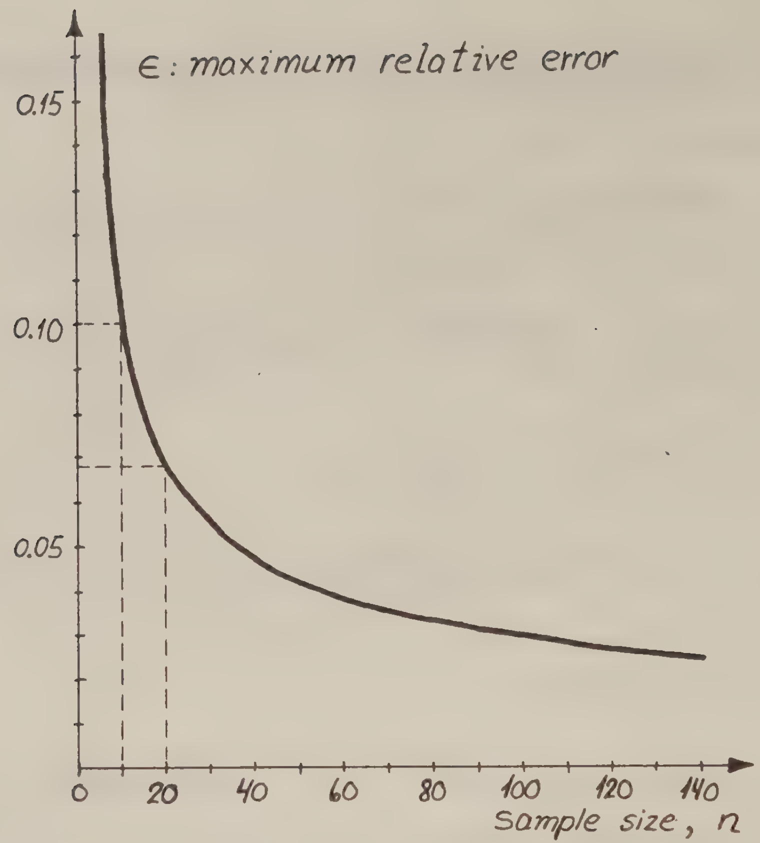


Figure 2.7.1 Maximum relative error (ϵ) of the estimate of the mean length as a function of sample size (n) (using data from Table 2.1.2)

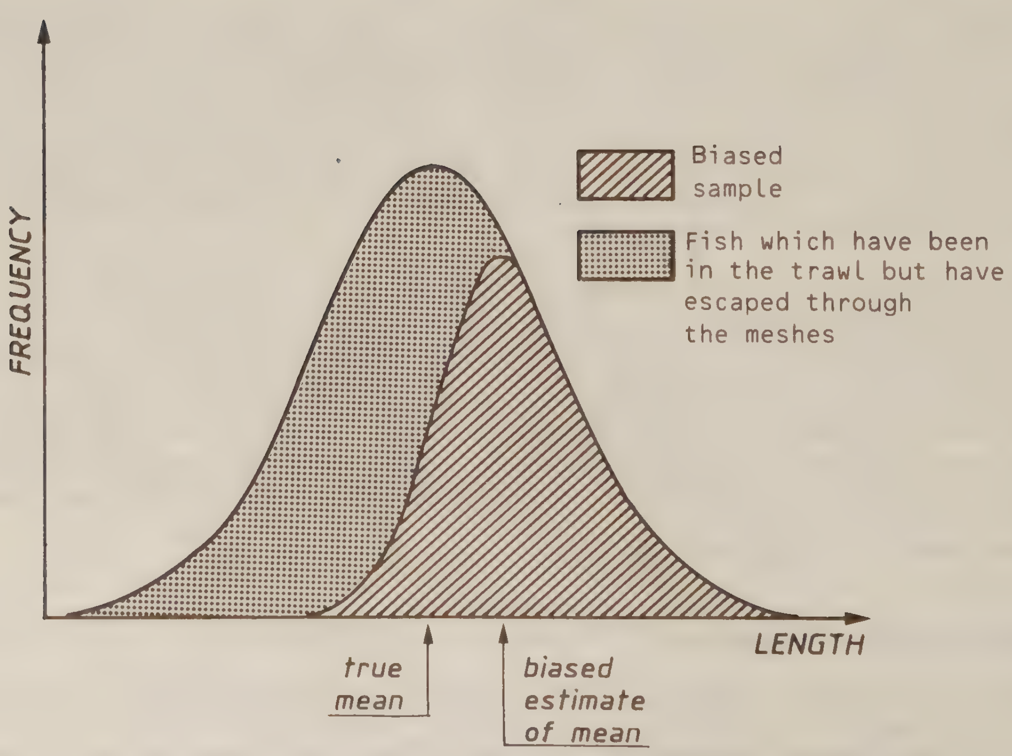

FREQUENCY
Biased sample
Fish which have been in the trawl but have escaped through the meshes
LENGTH
true mean
biased estimate of mean

Figure 2.7.2 Example of bias: bias caused by gear selection

estimate is "biased" if it deviates from the true value in a systematic manner. With an unbiased estimate we can approach the true value as closely as we want by increasing the sample size. With a biased estimate there will always be a deviation between the true value and the estimate, and the deviation is independent of the sample size.

To obtain an unbiased estimate of the mean length the sample should be a "random sample", i.e. any fish from the stock considered should have exactly the same probability of being sampled. Assuming that it is possible to obtain a random sample (this is usually very difficult in practice), how many fish, n, would we need in the sample to obtain a prespecified accuracy?

Suppose that we require an estimate of the mean length not deviating more than 7% from the true mean length and that we want to be 95% certain of this. We then would require the upper and lower 95% confidence limits not to deviate more than 7% from the estimated mean. Using Eq. 2.3.1 we get:

$$100\,\frac{t(n-1)\; s/\sqrt{n}}{\bar{x}} = 7\% \quad \text{or} \quad \frac{t(n-1)\; s/\sqrt{n}}{\bar{x}} = 0.07$$

Solving this equation with respect to n gives:

$$\left[\frac{t(n-1) * s}{\epsilon * \bar{x}}\right]^2 = n \qquad (2.7.1)$$

where ϵ stands for "maximum relative error" (in this case 0.07). In order to apply Eq. 2.7.1 for estimating the required sample size we must already know the standard deviation, s, from previous samples.

We can also solve Eq. 2.7.1 with respect to e, using, for example, the estimates of s = 2.20 and $\bar{x}$ = 15.07 from Table 2.1.2.

$$\epsilon = \frac{t(n-1) * 2.20}{15.07 * \sqrt{n}} = 0.146\,\frac{t(n-1)}{\sqrt{n}}$$

In Figure 2.7.1 the maximum relative error, ϵ, is shown as a function of the sample size, n. Note that we gain relatively little by increasing n when n > 50. An increase from n = 10 to n = 20, however, produces a reduction in relative error from 10.0% to 6.8%.

If unbiased random samples can be obtained, there is no problem in estimating the required sample size for any prespecified accuracy. However, samples are usually biased in one way or another. If small fish can escape through the meshes of a trawl we get an overestimate of the population mean length (see Fig. 2.7.2). This is one example of bias. If the larger fish can swim faster than the trawl is towed and thereby avoid being caught we have another type of bias.

However, once you are aware of the bias you can often adjust for it, and in the example of mesh selection you can try to estimate how many fish there would have been if all had been retained by the gear. This will be further discussed in Chapter 6.

Another type of bias occurs if the spatial distribution of the population depends on the size of the fish. For example, when juvenile fish concentrate in certain nursery grounds and gradually migrate to the fishing grounds this may introduce a bias, if this migration pattern is not well understood. Such a bias caused by migration will be discussed in Chapter 11.

Another example of possible bias you may encounter concerns the situation where it would be inefficient to measure the complete catch from a large haul. After sorting the catch by species you would take a subsample of, say, 100 fish of one species. If you selected the 100 fish one by one by simple "hand-picking" this would create a biased sample, since there is a tendency to select the larger specimens. A proper procedure would be to divide the catch of the species into boxes of approximately equal weight and then select some of these at random for the subsample.

So far we have considered a population consisting of a very large number of fish, so that the sample formed an insignificant part of the population. Some "populations" in fisheries biology, however, consist of a small number of units, so that a sample constitutes a significant part of it. This is the case when the sampled "population" is not fish, but for example the landing places in a certain area.

Suppose the purpose of a sampling scheme is to estimate the mean landings per landing place during a certain period, say, a month, and suppose that the total number of landing places is 100. If we sampled all 100 landing places then the variance of our estimate of the mean landings would be zero. In that case we simply have the true population mean. If we have personnel for sampling only 50 landing places there would be a certain variance in our estimate. However, this variance would not be s^2 (as defined by Eq. 2.1.2), since only 50% of the population is left to produce the variance, as distinct from a length frequency sample where practically 100% of the population would still be left.

We allow for this by applying the so-called "finite population correction factor":

$$(1 - n/N)$$

where N is the population size (N = 100 in the example) and n the sample size. Let Y(i) be the landings at landing place no. i of the sample, i=1,2,...n, and let $\bar{Y}$ be the mean landings of all landing places, then the estimate of the mean landings is:

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^{n} Y(i)$$

The variance of the estimate then becomes (cf. Eq. 2.3.2 and 2.1.2):

$$\mathrm{VAR}(\bar{Y}) = s^2 \frac{1}{n} (1 - n/N) \qquad (2.7.2)$$

where

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Y(i) - \bar{Y})^2 \qquad (2.7.3)$$

The confidence interval of $\bar{Y}$ is (cf. Eq. 2.3.1):

$$\left[\bar{Y} - t(n-1)\, s \sqrt{(1/n - 1/N)}\ ,\ \bar{Y} + t(n-1)\, s \sqrt{(1/n - 1/N)} \right]$$

The estimate of the total landings is:

$$Y = N\,\bar{Y} \qquad (2.7.4)$$

and its variance is

$$VAR(Y) = N^2 \; VAR(\bar{Y}) \tag{2.7.5}$$

Y is usually the quantity we are interested in. Eq. 2.7.5 follows from the general rule for a random variable, Eq. 2.3.3, where N is a constant.

Note that Eq. 2.7.2 is general in the sense that it also applies to large (in practice infinite) populations as the finite population correction factor, (1-n/N), becomes 1.0 when N is infinite. Often Eq. 2.7.4 is expressed as:

$$Y = (N/n) \sum_{i=1}^{n} Y(i) \;)$$

and we say that the sample has been raised to the total, Y, by application of the "raising factor" (N/n).

2.8 Stratified random sampling

Consider again the problem of estimating the total landings from the 100 landing places during a certain month, as dealt with in Section 2.7. Assume that a sampling programme has been conducted in the previous years, based on which the 100 landing places have been divided into three categories as shown in Table 2.8.1. Table 2.8.2 shows a numerical example (from Gulland, 1966) corresponding to Table 2.8.1. To obtain an estimate of s(j) a survey covering all landing places was carried out during one month. Such a division of the total population is called a "stratification" and the categories (large, medium, small) are called "strata".

The standard deviation, s(j), in Tables 2.8.1 and 2.8.2 is defined by the corresponding variance:

$$s(j)^2 = \frac{1}{N(j)-1} \sum_{i=1}^{N(j)} (Y(j,i) - \bar{Y}(j))^2 \tag{2.8.1}$$

where the stratum mean is:

$$\bar{Y}(j) = \frac{1}{N(j)} \sum_{i=1}^{N(j)} Y(j,i) \tag{2.8.2}$$

Note that Eqs. 2.8.1 and 2.8.2 produce the true parameters for the particular month in which the data were collected.

Often we are not in a position to sample all landing places. Let us assume that we have manpower, funds, etc. available only for a sample of n ($n < 100$) landing places. The sample size can be written as the sum:

$$n = n(1) + n(2) + n(3)$$

where n(j) is the number of landing places sampled in stratum no. j. To "design a sampling scheme" means to decide which sample sizes, n(1), n(2) and n(3), should be applied to each of the three strata.

One could now ask why we should complicate the sampling by the introduction of strata. The answer is that we (nearly) always obtain a more precise estimate of the population mean from stratified sampling than from simple random sampling. How much we gain depends on the choice of strata. If the observations within a stratum are approximately of the same size (as is the case in Table 2.8.2) we are likely to gain from the stratification. The less variation there is within a stratum, the more is gained by stratification. On the other hand, for practical reasons there is a limit to the number of strata that can be selected.

The basic ideas behind stratified sampling can be summarized as follows, the sample size, n(j), should be large when:

1) The stratum is large (if N(j) is large)
2) The standard deviation s(j) is large

Mathematically this can be expressed as

$$n(j) \text{ proportional to } N(j)\, s(j) \tag{2.8.3}$$

or

$$n(j) = n \frac{N(j)\, s(j)}{\Sigma\, N(j)\, s(j)} \tag{2.8.4}$$

where n is the total sample size. This formula is called the "optimum stratified sampling equation" (or "Neyman allocation").

Table 2.8.1 **An example of stratification (from Gulland, 1966)**

	landing place category or stratum	number of landing places by stratum	average landings by stratum	standard dev. within each stratum s(j)
1	large	N(1)	$\bar{Y}(1)$	s(1)
2	medium	N(2)	$\bar{Y}(2)$	s(2)
3	small	N(3)	$\bar{Y}(3)$	s(3)

Table 2.8.2 **Numerical example of a stratification based on samples from one month (from Gulland, 1966)**

Y(j,i) = landings at landing place no. i in stratum j.										
LARGE LANDING PLACES:	Y(1,i)									
N(1) = 10	45	59	87	41	71	25	9	69	10	7
$\bar{Y}(1)$ = 42.3										
s(1) = 28.91										
MEDIUM LANDING PLACES:	Y(2,i)									
N(2) = 30	17	13	19	26	1	8	27	11	12	26
$\bar{Y}(2)$ = 15.2	5	8	10	16	16	4	16	16	13	29
s(2) = 8.57	14	25	29	27	20	25	2	7	3	12
SMALL LANDING PLACES:	Y(3,i)									
N(3) = 60	2	6	7	0	1	2	1	5	4	7
$\bar{Y}(3)$ = 4.2	8	9	3	2	5	4	2	0	2	8
s(3) = 2.81	5	3	8	9	8	9	1	6	5	3
	3	4	7	5	5	3	2	4	6	1
	6	2	5	1	0	3	8	0	4	3
	3	5	5	0	7	0	9	7	9	0

Example A:

Let us assume that we have funds available to collect 100 observations, that there are two strata and that the strata sizes and their standard deviations are:

	Stratum 1	Stratum 2
N(j)	1000	2000
s(j)	50	10

Then, the 100 observations should be allocated to the two strata according to Eq. 2.8.4 in the following proportions:

$$\text{Stratum 1:} \quad \frac{1000 * 50}{1000 * 50 + 2000 * 10} = 0.714$$

$$\text{Stratum 2:} \quad \frac{2000 * 10}{1000 * 50 + 2000 * 10} = 0.286$$

The corresponding sample sizes then become:

Stratum 1: 100 * 0.714 = 71 observations

Stratum 2: 100 * 0.286 = 29 observations

Usually the budgets available for execution of sampling programmes are limited. The cost of taking a sample differs from stratum to stratum and it is therefore possible also to take the cost of sampling a unit (an observation) into account, when designing a sampling scheme.

Let c(j) be the cost of taking a sample unit in stratum j and let c(0) be the fixed cost of the sampling programme. The total cost for m strata then becomes:

$$C = c(0) + \sum_{j=1}^{m} c(j) * n(j) \qquad (2.8.5)$$

It can be shown that for an optimum allocation the sample size should be proportional to $1/\sqrt{c(j)}$. The rules for optimum allocation can now be extended with a third point, the sample size, n(j), should be large if:

3) Sampling is inexpensive

An optimum use of the available resources is then achieved when the sample size is proportional to

$N(j)\ s(j)/\sqrt{c(j)}$, or if

$$n(j) = n \frac{N(j)\ s(j)/\sqrt{c(j)}}{\Sigma\ N(j)\ s(j)/\sqrt{c(j)}} \qquad (2.8.6)$$

where n is the total sample size. Using the criterion Eq. 2.8.6, we minimize the variance of the estimate of the total, Y.

If the total budget available is C and the fixed cost is c(0) then the total number of sample units in all strata is given by:

$$n = (C - c(0)) \frac{\Sigma\ N(j)\ s(j)/\sqrt{c(j)}}{\Sigma\ N(j)\ s(j) * \sqrt{c(j)}} \qquad (2.8.7)$$

and the number of units sampled in stratum j is given by:

$$n(j) = (C - c(0)) \frac{N(j)\ s(j)/\sqrt{c(j)}}{\Sigma\ N(j)\ s(j) * \sqrt{c(j)}} \qquad (2.8.8)$$

Example B:

Let the total budget for the sampling programme given in Example A be 1800 (money units) while the fixed cost is 444 and the prices per sample unit are:

Stratum 1: 16
Stratum 2: 9

Then the total number of samples, n, that can be taken with the available budget is determined by Eq. 2.8.7:

$$(1800 - 444) * \frac{1000 * 50 / 4 + 2000 * 10 / 3}{1000 * 50 * 4 + 2000 * 10 * 3} =$$

$$= 1356 * 19167 / 260000 = 99.96, \text{ say 100 sample units}$$

These 100 sample units should then be allocated to stratum 1 and 2 respectively as follows:

$$n(1) = (1800 - 444) * \frac{1000 * 50 / 4}{1000 * 50 * 4 + 2000 * 10 * 3} = 65$$

$$n(2) = (1800 - 444) * \frac{2000 * 10 / 3}{1000 * 50 * 4 + 2000 * 10 * 3} = 35$$

Note that now more sample units are allocated to the inexpensive stratum 2, compared to Example A, where the price per sample unit was not taken into account (or assumed to be the same).

Hereby we conclude the theory for optimum allocation and turn to the question of how the estimates of mean values, the totals and their variances can be calculated. Although the theory given below is general, you may again think of the example given in Table 2.8.2. Having decided the sample sizes, we obtain an estimate of the mean landings of each stratum by:

$$\bar{Y}(j) = \frac{1}{n(j)} \sum_{i=1}^{n(j)} Y(j,i) \qquad (2.8.9)$$

Let m be the number of strata (j = 1,2,...,m), then the estimate of the total population mean is:

$$\bar{Y} = \sum_{j=1}^{m} \frac{N(j)}{N} \bar{Y}(j) \qquad (2.8.10)$$

and finally, the estimate of the total landings is

$$N\ \bar{Y}$$

Let the symbol "VARst" stand for the variance of an estimate obtained from stratified ("st") random sampling. The estimate of the variance of the total population mean, $\bar{Y}$ is:

$$\text{VARst}(\bar{Y}) = \sum_{j=1}^{m} (N(j)/N)^2 \, \text{VAR}(\bar{Y}(j)) \qquad (2.8.11)$$

Compare with Eq. 2.7.4 where the estimate of the variance of the estimate of the mean for stratum j, VAR($\bar{Y}(j)$), is defined by Eq. 2.7.2:

$$\text{VAR}(\bar{Y}(j)) = \frac{1-n(j)/N(j)}{n(j)} \, s(j)^2 \qquad (2.8.12)$$

Note the finite population correction: $1 - n(j)/N(j)$

The variance within stratum j is:

$$s(j)^2 = \frac{1}{n(j)-1} \sum_{j=1}^{n(j)} (Y(j,i) - \bar{Y}(j))^2 \qquad (2.8.13)$$

Eq. 2.8.11 follows from the general rules for the variance of a sum of independent random variables given en Eqs. 2.3.3 and 2.3.4.

Inserting Eq. 2.8.12 into Eq. 2.8.11 the latter may also be written:

$$\text{VARst}(\bar{Y}) = \frac{1}{N^2} \sum_{j=1}^{m} \frac{N(j)(N(j)-n(j))}{n(j)} \, s(j)^2 \qquad (2.8.14)$$

Eq. 2.8.14 is more convenient from a calculation point of view (cf. Exercise 2.7).

The variance of the total landings in stratum j is:

$$\text{VAR}(Y(j)) = N(j)^2 \, \text{VAR}(\bar{Y}(j)) \qquad (2.8.15)$$

and the variance of the total catch (all strata) is:

$$\text{VARst}(Y) = N^2 \, \text{VARst}(\bar{Y}) \qquad (2.8.16)$$

In the examples given above the strata were different sizes of landing places. There are usually a number of possible stratifications available, for example by:

Gear type
Boat type
Fishing season
Fishing ground
Species or species groups
Commercial size category

Usually one can take advantage of whatever stratification is available. Sometimes we are forced to stratify the sampling, as might be the case when fish are landed already sorted into commercial size categories. Sometimes we will have to do extra work to obtain the stratification. This may apply to a stratification by boat type. If we sample from an auction hall at the time the fish are sold, we may have to trace the boat which caught the sample. If we are to invest extra resources (time, funds and manpower) in obtaining a stratification the increased precision obtained from that stratification should be considered relative to the cost.

In some cases (e.g. when starting up) we may not know the variance within strata but only the stratum size, N(j). In that case it is recommended to use "proportional sampling", i.e. to allocate samples in proportion to the stratum size (cf. Exercise 2.7). Using proportional sampling the 100 observations of Example A above should be allocated to the two strata in the proportions:

$$\frac{1000}{1000 + 2000} = 0.333 \quad \text{and} \quad \frac{2000}{1000 + 2000} = 0.677$$

(Compare with the values obtained by optimum allocation, 0.714 for stratum 1 and 0.286 for stratum 2).

The corresponding sample sizes become (figures in brackets are the values in the case of optimum allocation):

Stratum 1: 100 * 0.333 = 33 (71) observations
Stratum 2: 100 * 0.667 = 67 (29) observations

Note that proportional sampling is only identical to optimum stratified sampling in the exceptional case that the variances of all strata are equal.

In Chapter 7 we shall discuss sampling of length frequency data from commercial catches (e.g. market sampling) and discuss the use of "raising factors". For further studies in the theory of sampling techniques see for example Raj (1968), Som (1973) or Cochran (1977).

(See **Exercise(s)** in Chapter 17).

3 ESTIMATION OF GROWTH PARAMETERS

The study of growth means basically the determination of the body size as a function of age. Therefore all stock assessment methods work essentially with age composition data. In temperate waters such data can usually be obtained through the counting of year rings on hard parts such as scales and otoliths. Such rings are formed due to strong fluctuations in environmental conditions from summer to winter and vice versa.

In the tropical areas such drastic changes do not occur and it is therefore very difficult, if not impossible to use this kind of seasonal rings for age determination. Only recently methods have been developed to use much finer structures, so called daily rings, to count the age of the fish in number of days. These methods, however, require special expensive equipment and a lot of manpower and it is therefore not likely, that they will be applied on a routine basis in many places.

Fortunately several numerical methods have been developed which allow the conversion of length frequency data into age composition. Although these methods do not require the reading of rings on hard parts, the final interpretation of the results becomes much more reliable if at least some direct age readings are available. The best compromise for stock assessment of tropical species is therefore an analysis of a large number of length frequency data combined with a small number age readings on the basis of daily rings. This manual does not deal with the techniques of age reading but references to special publications are given (see Section 3.2.1).

3.1 The von Bertalanffy growth equation

Pütter (1920) developed a growth model which can be considered the base of most other models on growth including the one developed as a mathematical model for individual growth by von Bertalanffy (1934), and which has been shown to conform to the observed growth of most fish species. The theory behind various growth models is reviewed by for example Beverton and Holt (1957), Ursin (1968), Ricker (1975), Gulland (1983), Pauly (1984) and Pauly and Morgan (1987), but we shall deal here only with the von Bertalanffy growth model of body length as a function of age. It has become one of the cornerstones in fishery biology because it us used as a submodel in more complex models describing the dynamics of fish populations. Fig. 3.1.0.1 illustrates the model in graphical as well as in mathematical form.

The mathematical model, B, expresses the length L as a function of the age of the fish, t:

$$L(t) = L_\infty [1 - \exp(-K(t-t_o))] \qquad (3.1.0.1)$$

The rigth hand side of the equation contains the age, t, and some parameters. They are: "L_∞" (read "L-infinity"), "K" and "t_o" (read "t-zero"). Different growth curves will be created for each different set of parameters, therefore it is possible to use the same basic model to describe the growth of different species, simply by using a special set of parameters for each species.

To illustrate the use of the model, assume that the three parameters have been estimated for some particular fish stock and that the values are:

L_∞ = 50 cm, K = 0.5 per year and t_o = -0.2 year

Then we insert these parameter values into the von Bertalanffy growth equation:

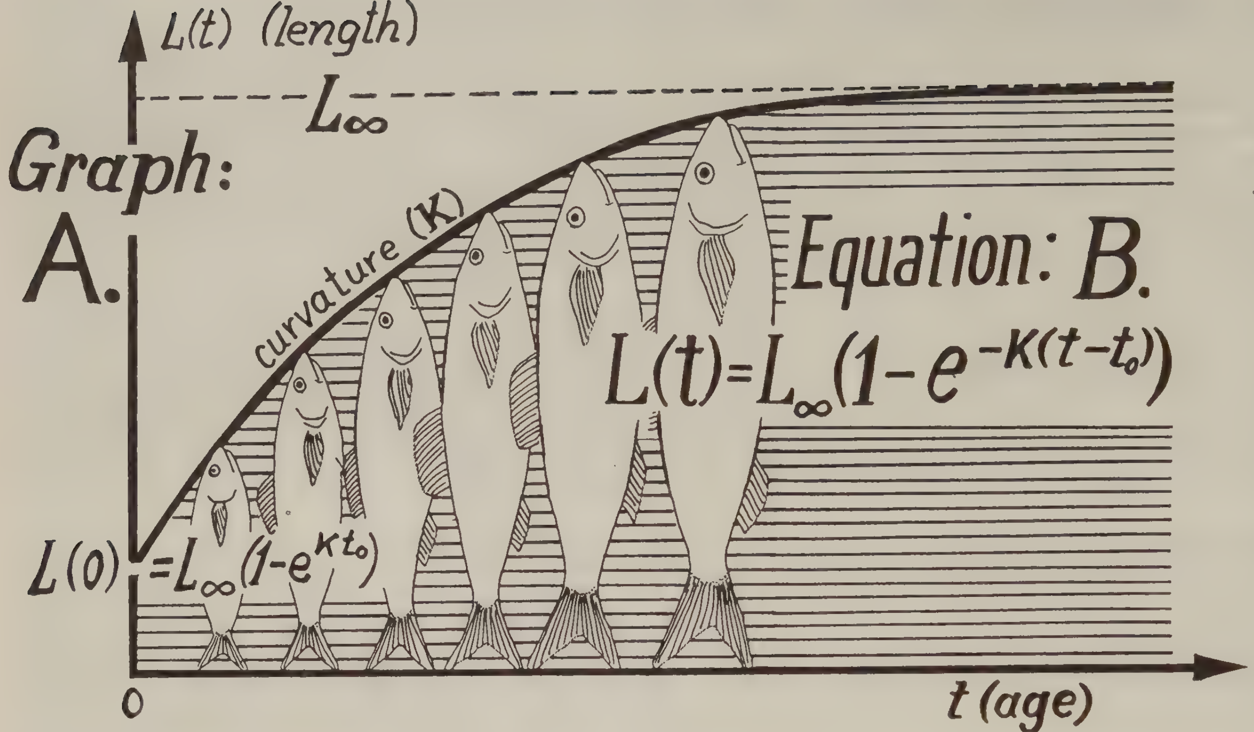

L(t) (length)
L_∞
Graph:
A.
curvature (K)
Equation: B.
$L(t)=L_\infty(1-e^{-K(t-t_0)})$
$L(0)=L_\infty(1-e^{Kt_0})$
0
t (age)

Fig. 3.1.0.1 The von Bertalanffy growth equation

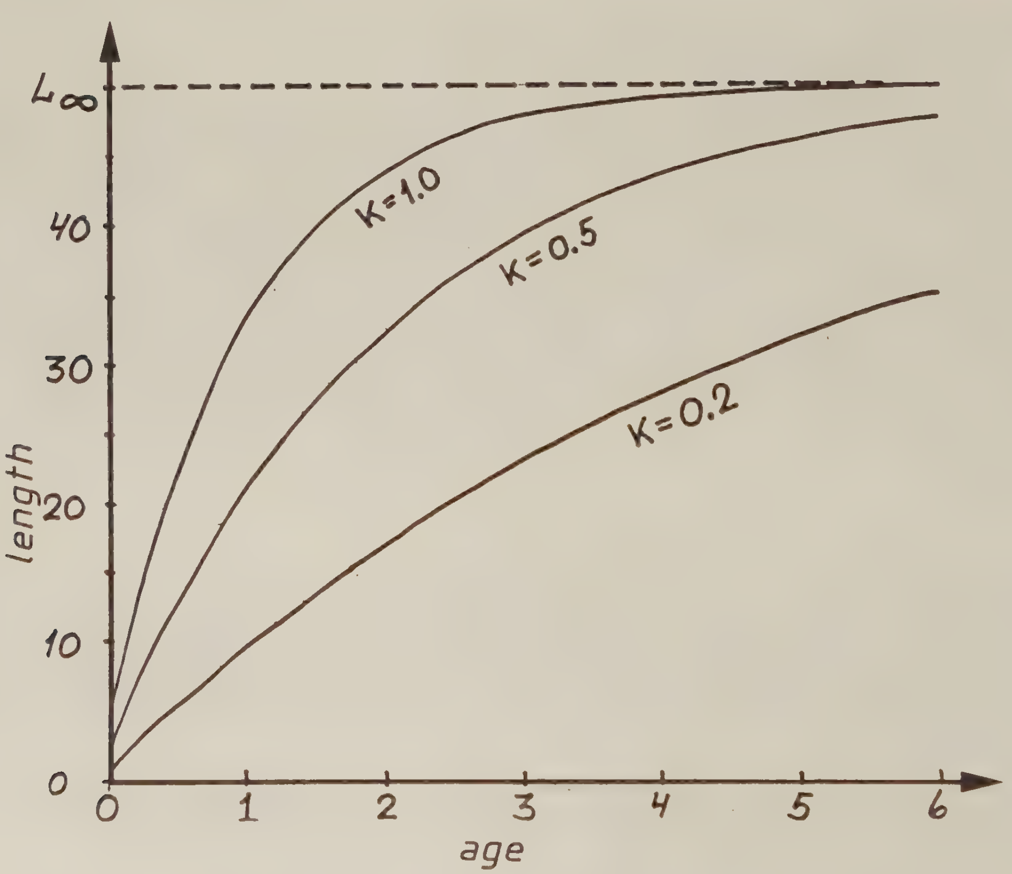


Fig. 3.1.0.2 A family of growth curves with different curvature parameters, different K values

$$L(t) = L_\infty [1 - \exp(-K * (t-t_o))] \qquad (3.1.0.1)$$

$$L(t) = 50 * [1 - \exp(-0.5 * (t + 0.2))]$$

The length in cm at a given age of an average fish of the stock in question can now be calculated by inserting a value for t, the age. For example, t = 2 years:

$$L(2) = 50 * [1 - \exp(-0.5 * (2 + 0.2))] = 33.4 \text{ cm}$$

Thus, knowing the parameters we can calculate the length at any age of the fish in the stock in question:

age of fish (year)	body length of fish (cm)
0.5	14.8
1.0	22.6
1.5	28.6
2.0	33.4
3.0	39.9
5.0	46.3
.... etc.	

From such a table a graph (growth curve) can be produced for this set of paramters, as in Fig. 3.1.0.1.

The parameters can to some extent be interpreted biologically. L_∞ is interpreted as "the mean length of very old (strictly: infinitely old) fish" (see Fig. 3.1.0.1). K is a "curvature parameter" which determines how fast the fish approaches L_∞. Some species, most of them short-lived, almost reach L_∞ in a year or two and have a high value of K. Other species have a flat growth curve with a low K-value and need many years to reach anything like L_∞. This is illustrated in Fig. 3.1.0.2. The third parameter, t_o, sometimes called "the initial condition parameter", determines the point in time when the fish has zero length. Biologically, this has no meaning, because the growth begins at hatching. The latter may be called L(0) when we put t = 0 at the day of birth. It is easily identified by inserting t = 0 into Eq. 3.0.1.0:

$$L(0) = L_\infty(1 - \exp(Kt_o)$$

However, L(0) may not be a realistic estimate of the length at birth because fish larvae do not always grow according to the von Bertalanffy model. The important point is that fish old enough to be exploited usually do.

Let us therefore turn the attention to the description of the growth of larger (exploited) fish. Not surprisingly, the rate of increase of the curve in Fig. 3.1.0.1 diminishes as the fish grows older. If L(t) is the length at age t then the growth rate per year

$$\frac{L(t+1.0) - L(t)}{1.0}$$

approaches zero when the fish grow very old. For example, column C of Table 3.1.0.1 shows the growth rates as a function of age for a curve as shown in Fig. 3.1.0.1. The concept of growth rate may be based on any other time period than a year, say, a day or a month. In general we consider a time period of duration Δt and define the "growth rate" by:

Table 3.1.0.1 Growth rate as a function of age corresponding to the growth curve in Fig. 3.1.0.1

A	B	C	D
age t years	length L(t) cm	growth rate $\frac{L(t+1)-L(t)}{1}$ cm/year	mean length $\frac{L(t+1)+L(t)}{2} = \bar{L}(t)$ cm
1	25.7		
		10.3	30.9
2	36.0		
		6.9	39.5
3	42.9		
		4.6	45.2
4	47.5		
		3.2	49.1
5	50.7		
		2.1	51.8
6	52.8		
		1.4	53.5
7	54.2		

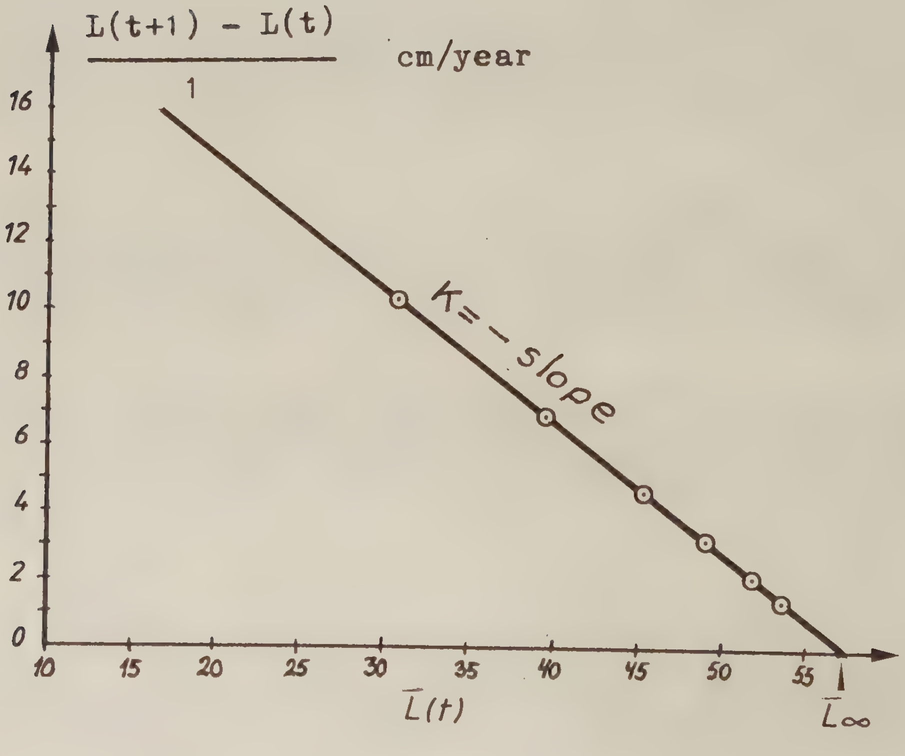


Fig. 3.1.0.3 Plot of growth rate against length. From columns C and D of Table 3.1.0.1

$$\frac{\Delta L}{\Delta t} = \frac{L(t+\Delta t) - L(t)}{\Delta t} \quad \text{cm/year} \qquad (3.1.0.2)$$

Time (or age) is often measured in units of years, but if we measured the growth rate per month, then

$$\Delta t = 1/12 \text{ years} = 0.0833 \text{ years}$$

or per day then,

$$\Delta t = 1/365 \text{ years} = 0.00274 \text{ years}$$

The simplest mathematical model which reflects the findings of column C of Table 3.1.0.1, i.e. that the growth rate decreases with time, is the linear function:

$$\frac{\Delta L}{\Delta t} = a + b\ L(t) \quad \text{cm/year} \qquad (3.1.0.3)$$

Apparently different from Eq. 3.1.0.1 this actually is just another way of writing the von Bertalanffy growth equation. We shall not here, however, concern ourselves with the mathematical proof.

It has turned out to be more convenient to write Eq. 3.1.0.3 in the form:

$$\frac{\Delta L}{\Delta t} = K\ (L_\infty - L(t)) \quad \text{cm/year} \qquad (3.1.0.4)$$

where the contants a and b have been replaced by

$$K = -b \quad \text{and} \quad L_\infty = -\ a/b$$

The plot of $\Delta L/\Delta t$ on L is shown in Fig. 3.1.0.1. The mean length over the year is used as the independent variable, L:

$$\bar{L}(t) = \frac{L(t+1) + L(t)}{2}$$

as given in column D of Table 3.1.0.1. The advantage of using K and L_∞ instead of a and b is that as already mentioned they are meaningful from a biological point of view.

From Eq. 3.1.0.4 it follows that

$$\text{if } L(t) = L_\infty \quad \text{then } \Delta L/\Delta t = K(L_\infty - L_\infty) = 0$$

i.e. when the fish achieves length L_∞ growth stops and L_∞ is thus the maximum length of a fish.

3.1.1 Variability and applicability of growth parameters

Growth parameters, of course, differ from species to species, but they may also vary from stock to stock within the same species, i.e. growth parameters of a particular species may take different values in different parts of its range. Also successive cohorts may grow differently depending on environmental conditions. Further growth parameters often take different values for the two sexes. If there are pronounced differences between the sexes in their growth parameters, the input data should be separated by sex and values of K, L_∞ and t_o should be estimated for each sex separately.

Although the physiology of crustaceans is very different from that of fishes, their body growth appears also to conform to the von Bertalanffy growth model (see e.g. Garcia and Le Reste, 1981). An individual crustacean (i.e. a shrimp or lobster) does of course not conform to the von Bertalanffy model, but to some "stepwise curve", with each step accounting for a moult (as illustrated in Fig. 3.1.1.1). Because members of a cohort moult at different times, the average growth curve of a cohort of crustaceans becomes a smooth curve (dotted line). For further discussion on the modelling of population dynamics of crustaceans see, for example, Jamieson and Bourne (1986) and Caddy (1987).

(See **Exercise(s)** in Chapter 17).

3.2 Input data for the von Bertalanffy growth equation

There are several ways of obtaining input data for the methods used to derive the growth parameters L_∞, K and t_o. The methods may be categorized roughly into three groups:

1) Age reading and length measurements combined
    a) data from resource surveys with a research vessel
    b) data from samples taken from commercial catches

2) Length measurements only
    a) data from resource surveys with a research vessel
    b) data from samples taken from commercial catches

3) Mark-recapture (tagging) experiments, where two (or more) length measurements are obtained, viz. at the time of marking (usually on a research vessel) and the time of recovery (usually by the commercial fishery). This method is excellent from a theoretical point of view but difficult and costly to implement. We shall not go further into this method, except for Exercise 3.3.1. (See also Jones, 1977).

Below we shall first consider 1a) in Section 3.2.1 and then 2b) in Section 3.2.2.

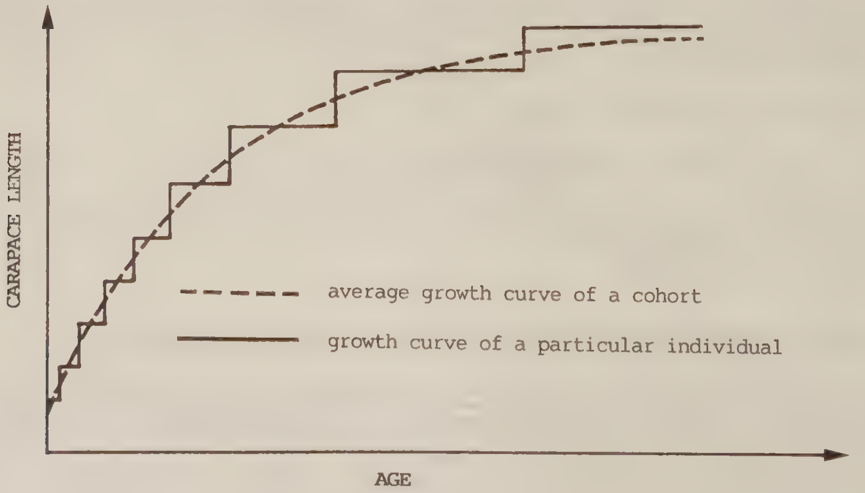


Fig. 3.1.1.1 Individual growth curve and average cohort growth curve of crustaceans

3.2.1 Data from age readings and length measurements

As has been stated in the introduction to this chapter, age reading is a relatively simple technique in the case of species from temperate waters, because their otoliths or scales show seasonal rings, one for the summer and one for the winter, which together form an annual ring. Sometimes such rings can be observed with the naked eye. In other cases simple techniques, such as burning can make them visible. The annual rings give sufficient information for most stock assessment purposes.

Unfortunately, tropical fish species seldom show clear annual rings in their otoliths or scales, because the strong seasonality which characterizes temperate zones is lacking. Recent discoveries, however, have created opportunities to also read ages of tropical fish, albeit within limited ranges and at a high cost in terms of manpower and initial investment. By going deeper into the formation of the rings in otoliths and scales it has been discovered that daily increments (or even increments caused by a certain food intake) can be detected by means of a strong microscope. The latest findings indicate that sometimes, the daily rings are so thin that they defeat the ordinary microscope whose detection power is limited by the wavelength of the light. Such rings can be read only by a scanning electron microscope (Morales-Nin, in press).

A large amount of literature has been produced on this subject in recent years, for example: Panella (1971), Bagenal (1974), Brothers (1980), Beamish and Mc Farlane (1983), Gjøsæter et al. (1983), Dayaratne and Gjøsæter (1986) and Williams (1986).

In a manual on tropical fish stock assessment it is necessary to concentrate on length measurements and consequently place less emphasis on age data. However, we are dealing with age here for two reasons. In the first place it may sometimes be possible to carry out a small number of age readings, which can be used to calibrate the findings obtained from length measurements alone. Secondly, it is easier to explain the concepts and theory on the basis of age and length data, than on the basis of length data only. Also we use data from research vessels to avoid further complications at this stage. The first example deals with data from a single survey, while the second example deals with data from a time series of surveys.

EXAMPLE 1: AGE/LENGTH COMPOSITION DATA FROM A SINGLE SURVEY

Suppose we have a random sample of fish from the stock of one species. This sample was taken by a survey with a duration of, say, a fortnight, in which trawl hauls were made over the entire distribution area of the stock in such a manner that the pooled data from all hauls made up a random sample (cf. Section 2.7). Suppose that the survey took place in the autumn of 1983 and that pooled length frequency data were obtained as presented in the last column of Table 3.2.1.1 (and also in Fig. 3.2.2.1). Suppose also that we have observed two annual peak recruitment seasons and therefore have decided to define two cohorts per year:

Spring cohort: Fish recruited from January to June
Autumn cohort: Fish recruited from July to December

The pooled data from all hauls can then be split into cohorts after the ages have been read and the mean length and the mean age of each cohort can be determined. The length frequencies for each cohort are also presented in Table 3.2.1.1. The type of age and length data shown in Table 3.2.1.1 is called an "age/length key" (this concept will be further discussed in Section 3.3.5). For each of these cohorts it is now possible to determine the mean age and mean length as given in Table 3.2.1.2.

Table 3.2.1.1 Age/length composition (hypothetical example). Basic data for Table 3.2.1.2. The graph for the total length frequency is shown in Fig. 3.2.2.1 ("-" stands for a zero observation)

recruitment season length interval cm	spring 1983	autumn 1982	spring 1982	autumn 1981	spring 1981	autumn 1980	survey October 1983 total of all hauls
12-13	1	-	-	-	-	-	1
13-14	4	-	-	-	-	-	4
14-15	11	-	-	-	-	-	11
15-16	24	-	-	-	-	-	24
16-17	38	-	-	-	-	-	38
17-18	42	-	-	-	-	-	42
18-19	33	-	-	-	-	-	33
19-20	20	-	-	-	-	-	20
20-21	7	-	-	-	-	-	7
21-22	2	1	-	-	-	-	3
22-23	-	3	-	-	-	-	3
23-24	-	5	-	-	-	-	5
24-25	-	8	-	-	-	-	8
25-26	-	11	-	-	-	-	11
26-27	-	14	-	-	-	-	14
27-28	-	16	1	-	-	-	17
28-29	-	15	1	-	-	-	16
29-30	-	13	2	-	-	-	15
30-31	-	11	3	-	-	-	14
31-32	-	7	4	-	-	-	11
32-33	-	4	6	1	-	-	11
33-34	-	2	7	1	-	-	10
34-35	-	1	7	1	-	-	9
35-36	-	-	8	2	-	-	10
36-37	-	-	7	3	1	-	11
37-38	-	-	6	3	1	-	10
38-39	-	-	5	4	1	-	10
39-40	-	-	4	4	2	1	11
40-41	-	-	3	5	2	1	11
41-42	-	-	2	4	2	1	9
42-43	-	-	1	3	2	1	7
43-44	-	-	-	3	3	1	7
44-45	-	-	-	2	2	1	5
45-46	-	-	-	2	2	2	6
46-47	-	-	-	1	2	2	5
47-48	-	-	-	1	1	1	3
48-49	-	-	-	-	1	1	2
49-50	-	-	-	-	1	1	2
50-51	-	-	-	-	1	1	2
51-52	-	-	-	-	-	1	1
total	182	111	67	40	24	15	439
mean length	17.3	27.9	35.3	40.2	43.3	45.5	
std. dev.	1.7	2.7	3.4	3.6	3.8	3.6	

Please note that the data presented in Table 3.2.1.1 are "hypothetical" or "faked" data. They were actually computed from a set of growth parameters (which determine the mean lengths for each cohort) and a set of standard deviations for the length distribution of each cohort. The advantage of using such hypothetical data in the context of this manual is that the true parameters are known, which is not the case with data taken from a real stock. That puts us in a position to compare the results of various

Table 3.2.1.2 **Hypothetical example of age and length composition data of species A from one research survey (derived from the "raw data" in Table 3.2.1.1)**

COHORT recruitment year	season	number observed	mean age (year)	mean length (cm)
1983	spring	182	0.64	17.3
1982	autumn	111	1.16	27.9
1982	spring	67	1.65	35.3
1981	autumn	40	2.10	40.2
1981	spring	24	2.64	43.3
1980	autumn	15	3.21	45.5
	total	439		

methods of parameter estimation with the real values. The data presented in Tables 3.2.1.1 and 3.2.1.2 will also be used as examples in Sections 3.4.1 and 3.4.2.

If we assume that all six cohorts in Table 3.2.1.2 have the same growth parameters, we can estimate the common growth parameters. In other words, we can determine the growth parameters which produce the growth curve giving the best fit to the pairs of mean age and mean length data. Exactly how this can be done is explained in Section 3.3.4.

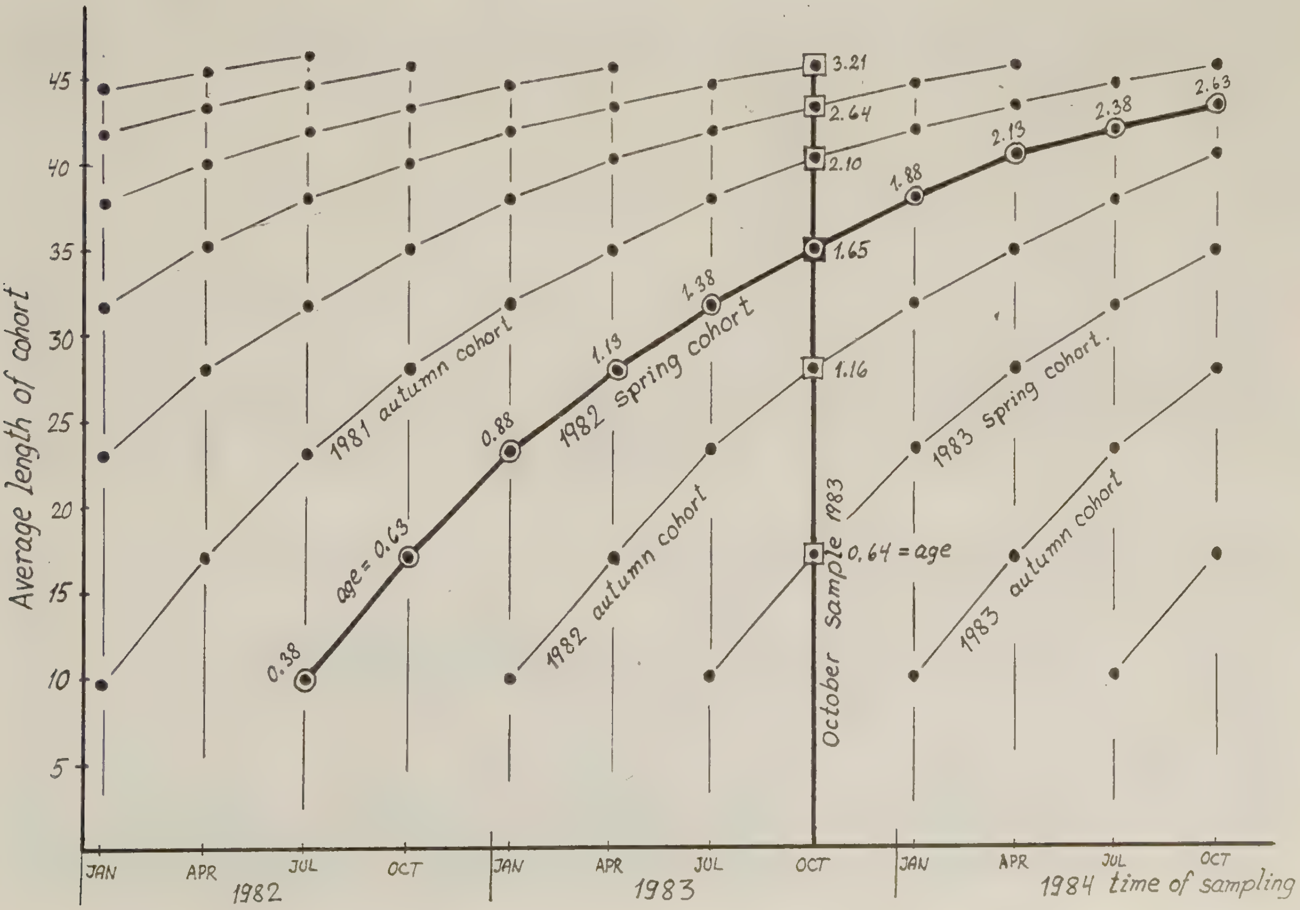


Fig. 3.2.1.1 Illustration of the data on age/length collected during a time series of surveys

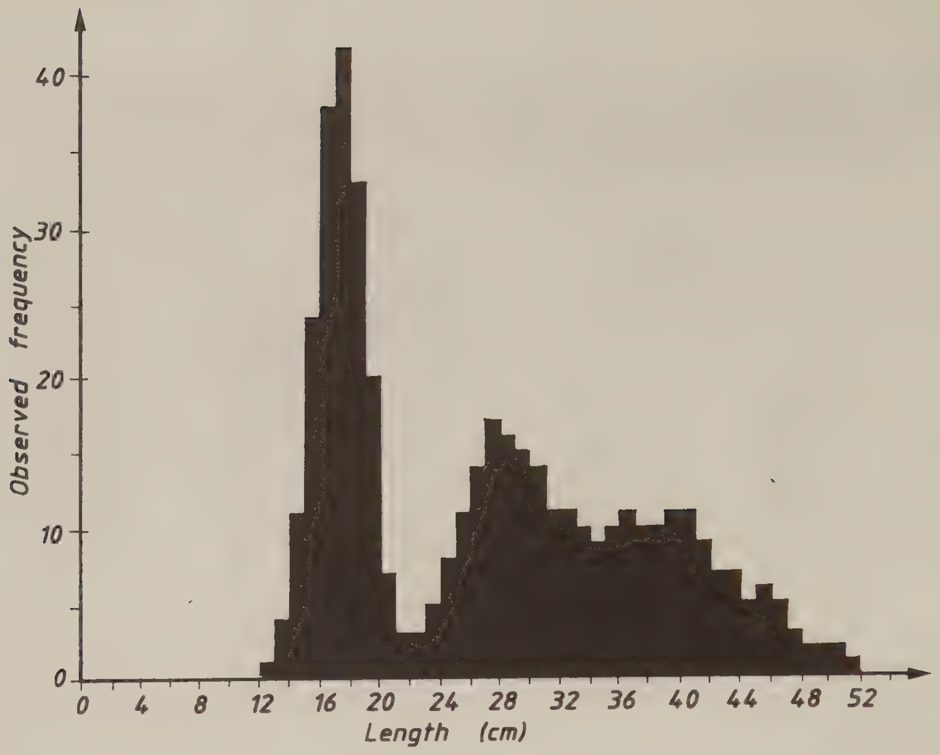


Fig. 3. 2. 2. 1.

Fig. 3.2.2.1 The length frequency sample, the only basic data in cases where age reading from hard parts is not possible. (Frequencies from the "total column" of Table 3.2.1.1)

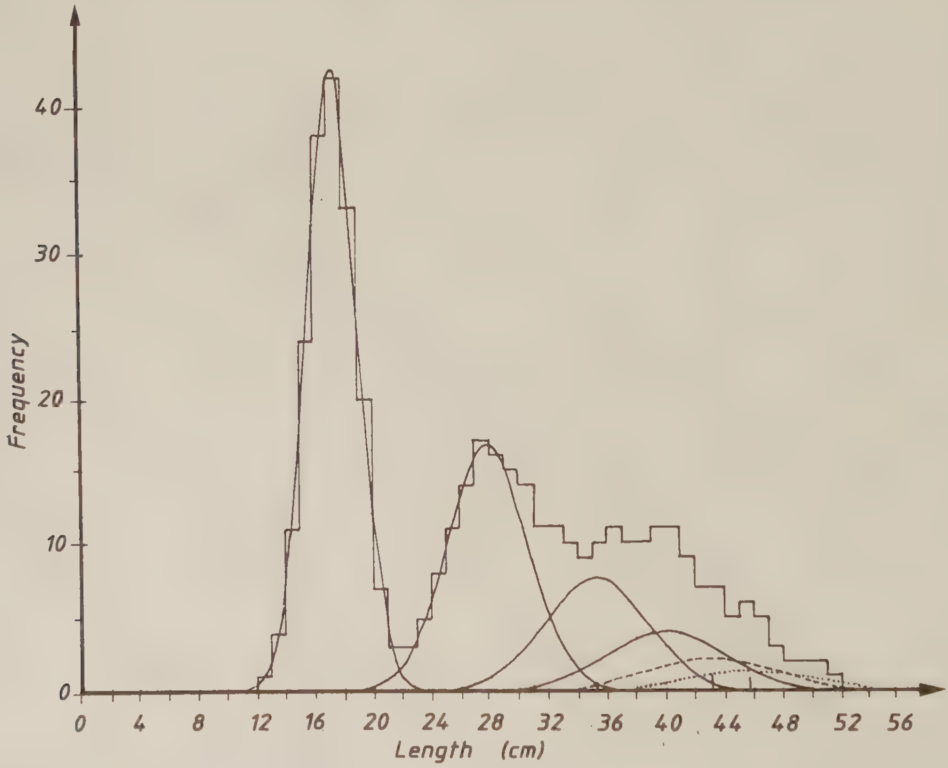


Fig. 3.2.2.2 The length frequency sample of Fig. 3.2.2.1, separated into normally distributed components. (Frequencies from the "total" column of Table 3.2.1.1). This example is also used to illustrate the "Bhattacharya method" described in Section 3.4.1 and the "maximum likelihood method" discussed in Section 3.5.3

EXAMPLE 2: AGE/LENGTH COMPOSTION DATA FORM MULTIPLE SURVEYS

If we now assume that the survey of Example 1 was only one of a series of 12 surveys carried out during the years 1982-1984, each year in the months of January, April, July and October, then such a survey programme would yield 12 tables like Table 3.2.1.1. By sampling the various cohorts regularly over a length of time, in this case three years, changes in the mean length can be determined by plotting them against the time of sampling as shown in Fig. 3.2.1.1. With this data set we are able to estimate the growth parameters for some of the cohorts individually. For the spring cohort in the recruitment year of 1982 for example, there are 10 pairs of age and length data, which can be used to estimate the parameters for that particular cohort.

The difference between following one particular cohort in time as shown here, and the determination of mean lengths of different cohorts at a certain moment, as presented in Example 1 is illustrated in Fig. 3.2.1.1, where the two different data types are indicated by heavy lines. The curve starting July 1982 and running to October 1984 shows the "real" growth of a cohort. The vertical line "October sample 1983" shows a "cross section" of the stock at that date.

In case of a short lived species (with a lifespan of, say, one to two years) we would have to follow a cohort in time as described in Example 2. The method based on a single sample would not be applicable because it would contain only one or two cohorts. Although there may be differences in the growth of different cohorts this difference is usually so small that it can be ignored. Data like those presented in Fig. 3.2.1.1 therefore could all be pooled into one data set and used in a way similar to the data of the October sample in 1983 (Table 3.2.1.1).

It is likely that bias will be less if sampling is done all the year round. Thus, although we can sometimes manage with a single sample, it is safer to use a time series of samples.

(See **Exercise(s)** in Chapter 17).

3.2.2 Length composition data (without age compositions)

Assume that we have a data set consisting of length frequencies of some species, but without age readings. The basic data set for a given sampling date would then look like the "total" (last) column of Table 3.2.1.1 or as drawn in Fig. 3.2.2.1. Is it possible to obtain a separation of the various cohorts which have contributed to this sample without using age reading techniques? The answer is that under certain conditions it is possible, except for parts where the ranges of length frequencies of different cohorts overlap each other too much.

The resolution of our sample into normally distributed components is shown in Fig. 3.2.2.2. These normally distributed components are approximately equal to the components given in the age/length key of Table 3.2.1.1. Because this is a hypothetical example there is no difficulty involved in the cohort separation. In practice, however, when the true cohort composition is unknown, the resolution into cohorts is not easy to obtain as will be demonstrated in Section 3.3. In particular the right hand part of the length frequency sample (i.e. the large fish) may create problems, which is evident in the case of Fig. 3.2.2.2. In Fig. 3.2.2.1 the youngest cohort, the spring cohort from 1983, can easily be distinguished by the naked eye. The next cohort (actually, the 1982 autumn cohort) is somewhat more difficult to see, while the remaining four cohorts may only be distinguished by using more sophisticated methods than simple visual inspection, or it may not be possible to separate them at all.

3.2.2.1 Commercial catches

Data for estimation of growth parameters may also be obtained from sampling the commercial catches. The basic principles in analysing samples from commercial landings are the same as for research data. The major difference lies in the bias problems. Commercial boats never attempt to collect a random sample of the stock, because they always go for the marketable sizes and try to find the areas with the highest concentrations of fish. However, keeping in mind sources of bias, and trying to stratify the sampling to minimize the bias, data from commercial fisheries may be used similarly to data from research surveys.

The major advantage of sampling commercial catches is that such samples are much cheaper to collect and thus sampling can be much more frequent than is possible in research surveys. In Chapter 7 problems in relation to sampling of commercial catches are further elaborated.

3.3 Estimation of growth parameters from length at age data

In this section we assume that pairs of observations of age and length are available. They may either be derived from readings of ring structures in hard parts or from length frequency analysis (Sections 3.4 and 3.5). Input data are either in the detailed from of an age/length composition as in Table 3.2.1.1 or in the processed form shown in Table 3.2.1.2. They may or may not be derived from a time series of samples (cf. Fig. 3.2.1.1). In the following we use for simplicity the input data format illustrated by Table 3.2.1.2.

Growth parameters can be derived from such data by graphical methods or plots, which are always based on a conversion to a linear equation, as discussed in Chapter 2. These plots are named after the authors of the papers wherein they were first described, viz. Gulland and Holt (1959), Ford-Walford (1933 and 1946 respectively) and von Bertalanffy (1934). Other methods that will be discussed are the "least squares method" and the use of age/length keys.

3.3.1 The Gulland and Holt plot

The Gulland and Holt (1959) plot was introduced in Section 3.1 by Eq. 3.1.0.4:

$$\Delta L/\Delta t = K\ L_\infty - K\ L(t) \qquad (3.3.1.1)$$

Using L(t) as the independent variable and $\Delta L/\Delta t$ as the dependent variable Eq. 3.3.1.1 becomes a linear regression:

$$\Delta L/\Delta t = a + b\ L(t)$$

The parameters K and L_∞ are obtained from:

$$K = -\ b \quad \text{and} \quad L_\infty = -\ a/b$$

This method does not produce an estimate of t_0. Table 3.1.0.1 and Fig. 3.1.0.3 show an example. In this case the length increment per year is plotted against the mean length during the corresponding year. Actual data are taken from columns C and D of Table 3.1.0.1.

The regression analysis gives:

$a = 22.47$ and $b = -\ 0.394$ from which we get

$K = -\ b = 0.4$ per year, and $L_\infty = -\ a/b = 57$ cm

Table 3.3.1.1 Input data for the Gulland and Holt plot and regression analysis (data derived from Table 3.2.1.1)

				Y	X
t	Δt	L(t)	$\Delta L(t)$	$\frac{\Delta L(t)}{\Delta t}$	$\frac{L(t+\Delta t) + L(t)}{2}$
0.64		17.3			
	0.52		10.6	20.4	22.6
1.16		27.9			
	0.49		7.4	15.1	31.6
1.65		35.3			
	0.45		4.9	10.9	37.7
2.10		40.2			
	0.54		3.1	5.7	41.8
2.64		43.3			
	0.57		2.2	3.9	44.4
3.21		45.5			

b (slope) = - 0.7670, a (intercept) = 38.52, n = 5

$$sb = \frac{1}{n-2}\left[(sy/sx)^2 - b^2\right] = \frac{1}{3}\left[(6.7727/8.7362)^2 - 0.7670^2\right] = 0.004216$$

sb = 0.06493, t(n-2) = 3.18

95 % confidence limits for b: [- 0.974 , -0.561] (cf. Section 2.4)

$$sa^2 = sb\left(\frac{n-1}{n} sx^2 + \bar{x}^2\right) = 0.004216\left(\frac{4}{5}\, 8.7362^2 + 35.62^2\right) = 5.606$$

sa = 2.37

95 % confidence limits for a: [31.0 , 46.1]

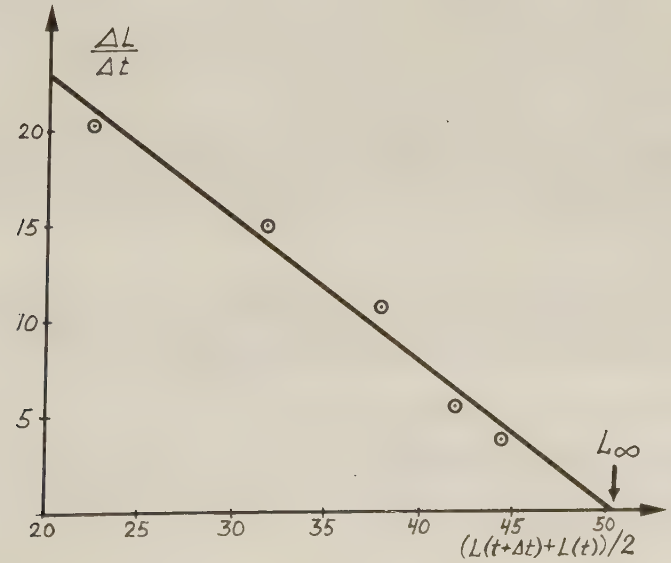


Fig. 3.3.1.1 Gulland & Holt plot corresponding to Table 3.3.1.1 (hypothetical example). The intersection point between the regression line and the L-axis gives L_∞

The length "L(t)" in Eq. 3.3.1.1 represents the length range from L(t) at age t to $L(t+\Delta t)$ at age $t + \Delta t$. Thus, the natural quantity to enter into Eq. 3.3.1.1 is the mean length (cf. the example in Table 3.1.0.1):

$$\bar{L}(t) = (L(t+\Delta t) + L(t))/2$$

Only if Δt is small $\bar{L}(t)$ may be a reasonable approximation to the mean length.

Another example of the Gulland and Holt plot can be derived from Table 3.2.1.2 as shown in Table 3.3.1.1. From the estimates of intercept and slope we get:

$$K = - b = 0.8 \text{ per year}, \quad L_\infty = - a/b = -38.103/-.757 = 50 \text{ cm}$$

The 95% confidence limits for K are the same as those for b but with the sign changed, i.e. [0.54 , 0.97] (cf. Table 3.3.1.1). The confidence limits of L_∞ are more complicated to obtain. However, we may say that if K = 0.8 then the confidence interval of L_∞ becomes: [30.3/0.8 , 45.9/0.8] = [37.9 , 57.4]. These are not the confidence limits of L_∞, but only the limits conditioned on K = 0.8. (Actually, the confidence limits of a ratio are not defined, but the theory dealing with this is considered outside the scope of the present manual). Fig. 3.3.1.1 shows the Gulland and Holt plot corresponding to Table 3.3.1.1.

In Section 3.1 it was stated that it can be proved mathematically that Eq. 3.1.0.4: $\Delta L/\Delta t = K(L_\infty - L(t))$ is equivalent to

Eq. 3.1.0.1: $L(t) = L_\infty [1-\exp(-K(t-t_0))]$.

This, however, is correct only if the time interval, Δt, is infinitesimal. Thus, the Gulland and Holt plot, which is based on Eq. 3.1.0.4, is an approximation which is reasonable only for small values of Δt.

It is, however, possible to turn the Gulland and Holt plot into an "Exact Gulland and Holt plot", i.e. a regression analysis, which is correct for any value of Δt, if the time interval remains constant, i.e. if we have pairs of observations:

$$(t,L(t)) , (t+\Delta t,L(t+\Delta t)) , (t+2\Delta t,L(t+2\Delta t)), \dots$$

where Δt remains constant. It can be shown that Eq. 3.1.0.1 implies that:

$$L(t+\Delta t) - L(t) = b * L_\infty - b * L(t) \qquad (3.3.1.2)$$

where $b = 1 - \exp(-\Delta t * K)$. Thus if Δt remains constant b will remain constant and Eq. 3.3.1.2 becomes a linear regression:

$$Y = a - b * L(t)$$

with $Y = L(t+\Delta t) - L(t)$, $a = b * L_\infty$ and $b = 1 - \exp(-\Delta t * K)$. Note that the slope is - b and also, that in Eq. 3.3.1.2 the abscissa has the smaller of the two lengths, not their mean value.

The growth parameters are derived from:

$$K = - (1/\Delta t) \ln(1-b) \text{ and } L_\infty = a/b$$

(See **Exercise(s)** in Chapter 17).

3.3.2 The Ford-Walford plot

This method introduced by Ford (1933) and Walford (1946) has gained wide application. From the original growth equation (Eq. 3.1.0.1) it follows from a series of algebraic manipulations that:

$$L(t + \Delta t) = a + b\, L(t) \qquad (3.3.2.1)$$

where $a = L_\infty (1 - b)$ and $b = \exp(- K\, \Delta t)$

The parameters K and L_∞ are obtained from:

$$K = -\frac{1}{\Delta t} \ln b \quad \text{and} \quad L_\infty = \frac{a}{1-b}$$

To illustrate the use of Eq. 3.3.2.1 consider the figures in column B of Table 3.1.0.1. These are lengths for a series of ages with time intervals of one year. In Table 3.3.2.1 columns A and B contain the pairs of consecutive lengths, $(L(t), L(t+\Delta t))$, from Table 3.1.0.1. Using these as (x,y) Eq. 3.3.2.1 can be used for linear regression. However, this is only the case if a and b are constants. As K and L_∞ are constants, a and b become constants if Δt is a constant, i.e. if the time difference between consecutive lengths remains constant.

Carrying out the regression analysis we get:

$a = 18.70$ and $b = 0.672$ from which $K = 0.4$ per year and $L_\infty = 57$ cm,

the results already given in Section 3.1.

The Ford-Walford plot corresponding to Table 3.3.2.1 columns A and B is shown in Fig. 3.3.2.1. L_∞ can be estimated graphically from the intersection point of the regression line and the line $L(t) = L(t+\Delta t)$ (the 45 degree line) because for very old fish, which have stopped growing we have $L_\infty = L(t) = L(t+\Delta t)$. Note that if one has data for a Ford-Walford plot it is also possible to do the exact Gulland and Holt plot.

Columns C and D in Table 3.2.2.1 show the input for the ordinary Gulland and Holt plot derived from columns A and B (recall that $\Delta t = 1$ year in this case). The regression of column C on column D gives:

$b = -.39$ ($K = -b = 0.4$ per year) and $a = 22.46$ ($L_\infty = -a/b = 57$ cm)

The same result was found by the Ford-Walford plot. The two methods may not give precisely the same results, unless the data conform exactly to the von Bertalanffy equation. The data of Table 3.3.2.1 were backcalculated from the von Bertalanffy equation, so they give the same result in both cases.

As Δt does not need to remain constant in the Gulland and Holt plot, it is more widely applicable than the Ford-Walford plot. The initial condition parameter, t_o, is not estimated by the Ford-Walford plot, as was the case for the Gulland and Holt plot.

A third method to estimate L_∞ is the "Wetherall et al. method", which estimates both the total mortality coefficient and L_∞ at the same time. Because the method also deals with mortalities it is presented in Section 4.5.

Table 3.3.2.1 Pairs of consecutive lengths, with /\t = 1 year.
A and B: Input data for the Ford-Walford plot.
C and D: Input data for the Gulland and Holt plot.
Data derived from Table 3.1.3.1 (compare Fig. 3.3.2.1)

	A	B	C	D
t	x L(t)	y L(t+/\t)	y /\ L//\t	x $\bar{L}$
1	25.7	36.0	10.3	30.9
2	36.0	42.9	6.9	39.5
3	42.9	47.5	4.6	45.2
4	47.5	50.7	3.2	49.1
5	50.7	52.8	2.1	51.8
6	52.8	54.2	1.4	53.5

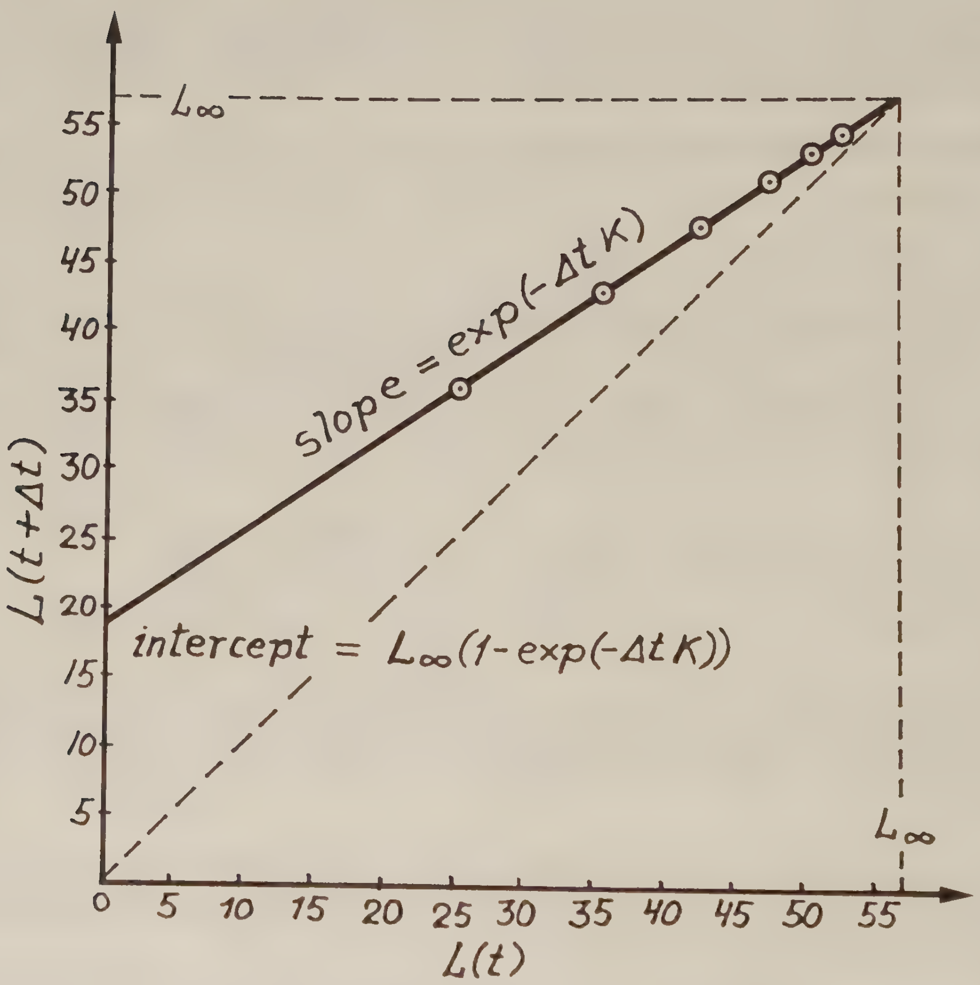


Fig. 3.3.2.1 Ford-Walford plot. Data from columns A and B of Table 3.3.2.1

3.3.3 The von Bertalanffy plot

The first method for estimating the von Bertalanffy growth parameters was suggested by von Bertalanffy (1934). It can be used to estimate K and t_o from age/length data, while it requires an estimate of L_∞ as input.

The growth equation (Eq. 3.1.4.1) can be rewritten:

$$- \ln(1 - L(t)/L_\infty) = - K * t_o + K * t \qquad (3.3.3.1)$$

With the age, t, as the independent variable (x) and the left-hand side as the dependent variable (y) the equation defines a linear regression, where K represents the slope and $- K * t_o$ the intercept.

Table 3.3.3.1 shows how to calculate the input data for the von Bertalanffy plot, based on data from Table 3.3.1.1 with $L_\infty = 50$ cm. The plot is shown in Fig. 3.3.3.1.

The von Bertalanffy plot is a more robust method than the Gulland and Holt plot (and the Ford-Walford plot) in the sense that it nearly always gives a reasonable estimate of K, given that a reasonable estimate of L_∞ is used in the computations (as illustrated in Exercise 3.3.3). One must ascertain, however, that the plot (Fig. 3.3.3.1) "looks" linear. On the other hand, one can say that the Gulland and Holt plot is stronger in the sense that it is better in bringing out cases where the observations are in conflict with the von Bertalanffy model.

Recalling the interpretation of L_∞ as the average length of a very old fish, there are various shortcut methods for estimating L_∞ for use in the von Bertalanffy plot:

1) In small samples you may simply use the largest fish.

2) In a very large sample you may take the average of the lengths of, say, the ten largest fish.

3) Perhaps the best way of estimating L_∞ is the Wetherall *et al.* method described in Section 4.5.

It may not matter as much as one might think which estimate of L_∞ is used. If you overestimate L_∞ the K will be underestimated, and together they will balance out, so that the resulting growth curve remains nearly the same for the range of ages represented in the data set. (This aspect will be further discussed in Section 3.4). There is, however, a problem in using the von Bertalanffy plot in connection with the definition of L_∞. The argument of the logarithm in Eq. 3.3.3.1, i.e. $(1 - L(t)/L_\infty)$ must be positive as the logarithm would otherwise not be defined.

Thus the von Bertalanffy plot cannot accept a length greater than L_∞, whereas with the definition of L_∞ as given in Section 3.1.4 it may well happen that for the very old fish, $L(t) > L_\infty$ because the observations (t, L(t)) fluctuate at random about the line. The von Bertalanffy plot actually uses the "inverse von Bertalanffy equation":

$$t(L) = t_o - \frac{1}{K} \ln(1 - L/L_\infty) \qquad (3.3.3.2)$$

which is Eq. 3.1.4.1 solved for t. It may be necessary to omit the oldest fish to obtain that $1 - L/L_\infty > 0$.

The L_∞ concept as applied in the von Bertalanffy plot is different from that applied in the Gulland and Holt plot, for the same reasons as the parameters in the "inverse linear regression" differ from those of the "ori-

Table 3.3.3.1 Input data and regression for the von Bertalanffy plot (data derived from Table 3.3.1.1, L_∞ = 50 cm)

x		y
t	L(t)	$-\ln(1 - L(t)/L_\infty)$
0.64	17.3	0.425
1.16	27.9	0.816
1.65	35.3	1.224
2.10	40.2	1.630
2.64	43.2	1.995
3.21	45.5	2.408

a = - 0.0658 b = K = 0.78
t_o = -a/b = 0.084

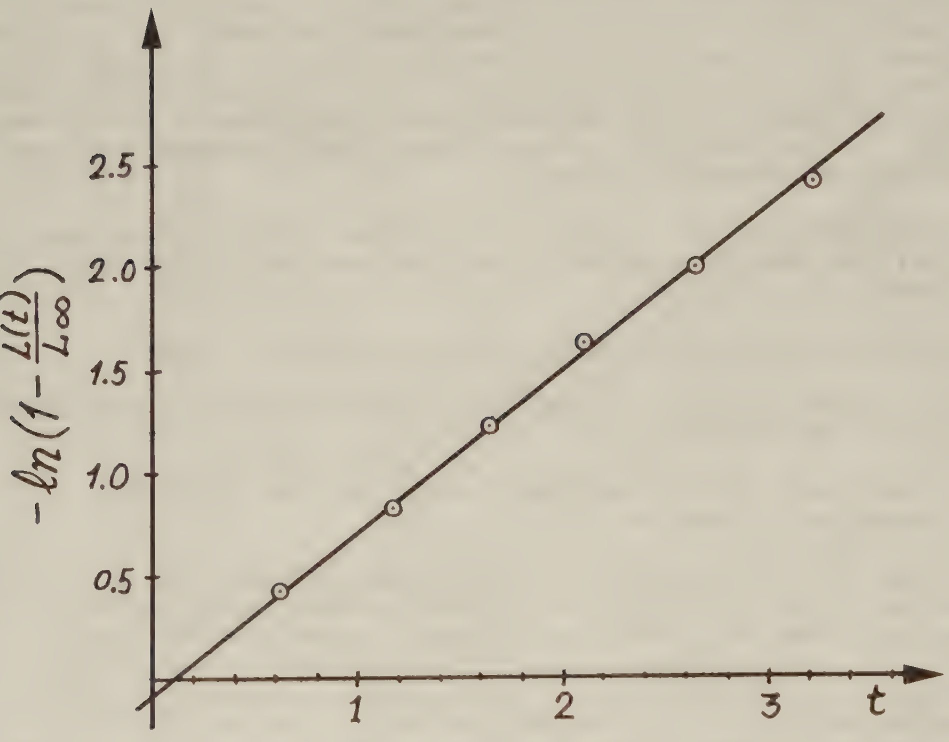


Fig. 3.3.3.1 Von Bertalanffy plot for the data in Table 3.3.3.1

Table 3.3.5.1 Hypothetical age/length key

length group cm	age group 0	age group I	age group II
5 - 10	1.0	0	0
10 - 15	0.25	0.75	0
15 - 20	0	0.5	0.5
20 - 25	0	0.1	0.9

ginal linear regression", unless the correlation coefficient is 1 or -1 (cf. Section 2.5).

(See **Exercise(s)** in Chapter 17).

3.3.4 The least squares method

This method is assumed to be superior to the methods introduced in the foregoing subsections from an estimation theory point of view. It is the non-linear parallel to the linear regression analysis introduced in Section 2.4. However, the computational work involved is considerable and in practice you need a computer to do the calculations.

Assume a series of pairs of obervations (length, age) to be available. These may have been obtained by age reading (cf. Section 3.2.1) or they may have been derived from modal progression analysis (to be discussed in Section 3.4.2). Let there be n observations:

(L(i),t(i)) = (length of fish no. i, age of fish no. i)

i = 1,2,....n.

The method estimates the growth parameters in such a way that the sum of the squares of the deviations between the model and the observations is minimized, i.e. it minimizes the sum with respect to the parameters L_∞, K and t_o:

$$\sum_{i=1}^{n} [L(i) - L_\infty * [1 - \exp(-K * (t(i)-t_o))]]^2 \qquad (3.3.4.1)$$

The LFSA package of microcomputer programs for fish stock assessment (Sparre, 1987) contains the program "VONBER" which can do the least square estimation of growth parameters. Many other similar computer programs are available (see Chapter 14)

The method used by this program is rather complicated and a full explanation falls outside the scope of this manual. However, conceptually, non-linear regression analysis is not more complex than linear regression, just as the square root of 3 is conceptually not more complex than the square root of 4, but the latter is much easier to calculate.

3.3.5 The use of age/length keys

An age/length key is a table showing, for each length class of fish, the percentage or fractional age frequency distribution, see Table 3.3.5.2. It is used in cases when an age determination of the fish is possible, but time consuming and therefore expensive. After the key has been prepared a large number of fish which were only measured can be distributed on age groups according to the key.

The age/length key of Table 3.3.5.2 could be based on 182 randomly drawn fish with the following length distribution:

Length group (cm)	5-10	10-15	15-20	20-25	Total
Frequency	110	40	22	10	182

The next step is to age the fish in each length group. Let us assume that the results are as shown in Table 3.3.5.1. Table 3.3.5.2 is then derived from Table 3.3.5.1 simply by dividing each row entry by the row total for each length group.

Table 3.3.5.2 Input data for estimation of an age/length key (hypothetical example)

length group cm	age group 0	age group I	age group II	total
5 - 10	110	0	0	110
10 - 15	10	30	0	40
15 - 20	0	11	11	22
20 - 25	0	1	9	10
total	120	42	20	182

Table 3.3.5.3 Age composition of a large length-frequency sample estimated by use of the age/length key in Table 3.3.5.1.

length group cm	age group 0	age group I	age group II	total
5-10	12088	0	0	12088
10-15	1759	5276	0	7035
15-20	0	894	894	1788
20-25	0	13	117	130
total	13847	6183	1011	21041

Table 3.3.5.4 Age/length key for *Scomberomorus brasiliensis* based on otolith readings, in percentages per 3-cm length groups. From Sturm, 1974

length group cm	age group (year)									
	0	I	II	III	IV	V	VI	VII	VIII	IX
13-16	100	-	-	-	-	-	-	-	-	-
16-19	-	100	-	-	-	-	-	-	-	-
19-22	-	100	-	-	-	-	-	-	-	-
22-25	-	100	-	-	-	-	-	-	-	-
25-28	-	96	4	-	-	-	-	-	-	-
28-31	-	55	45	-	-	-	-	-	-	-
31-34	-	5	95	-	-	-	-	-	-	-
34-37	-	-	91	9	-	-	-	-	-	-
37-40	-	-	73	27	-	-	-	-	-	-
40-43	-	-	33	63	2	-	-	-	-	-
43-46	-	-	15	77	8	-	-	-	-	-
46-49	-	-	5	65	29	-	-	-	-	-
49-52	-	-	1	47	50	2	-	-	-	-
52-55	-	-	-	38	51	11	-	-	-	-
55-58	-	-	-	10	62	21	7	-	-	-
58-61	-	-	-	3	50	25	22	-	-	-
61-64	-	-	-	-	19	44	31	6	-	-
64-67	-	-	-	-	-	66	17	17	-	-
67-70	-	-	-	-	-	-	75	25	-	-
70-73	-	-	-	-	-	-	-	33	33	33
> 73	-	-	-	-	-	-	-	-	50	50

Table 3.3.5.2 can be used to convert a much larger length frequency sample of the same stock (for which the age composition is unknown) into age groups. As an example we consider the length frequency sample of 21041 fish (which may or may not be a random sample):

Length group (cm)	5-10	10-15	15-20	20-25	Total
Frequency	12088	7035	178	8130	21041

By distributing the numbers in each length group upon the age groups according to the proportions given in Table 3.3.5.2, we get the results presented in Table 3.3.5.3. Length group 10-15 cm, for example, is estimated to consist of 7035 * 0.25 = 1759 0-group fish and 7035 * 0.75 = 5276 I-group fish. Adding the column entries we finally arrive at the age composition given in the bottom row of Table 3.3.5.3.

In order to estimate the age composition of the catch from a particular stock we need to determine an age/length key based on a small sample of age readings and reduce further sampling to the collection of length frequency data. Afterwards, lengths are converted to ages by means of the key. The same key may be used in consecutive years as long as there is no suspicion of major changes in the stock composition. However, in a period of for instance, markedly increasing effort the old fish may be disappearing from the catches. Table 3.3.5.4 shows an example of an age/length key for the Spanish mackerel Scomberomorus brasiliensis, a longlived tropical fish. To illustrate the limitations of an age/length key, consider the percentage age distribution of fish of 61-64 cm long, which are 4-7 years old. Now, if the fishing mortality (effort) increases markedly, most fish older than 5 years may be exterminated and the few fish of 61-64 cm still caught will be fastgrowing 4-year-olds with a few remaining 5-year-olds thrown in. Using the key on the new length frequency distributions would give the impression that the 61-64 cm fish are still 4-7 years old, with the 5-group dominating.

When collecting samples for an age/length key it is important to include in the sample some very small fish, and some very large ones. Otherwise, when large numbers of length measurements are to be distributed on age groups it will be found that some size classes represented are not in the key at all. When small and large fish are deliberately overrepresented in the key it is important to remember, that the key data alone cannot be used for estimation of growth or mortality parameters.

The methodology of fish stock assessment can in fact be entirely based on age/length compositions alone. The application of mathematical growth models is not necessary. To a certain degree this is the case in the assessments made in the ICES area (Northeast Atlantic). However, since reliable age/length keys are not likely to become available for most tropical species in the near future, as well as for a number of other reasons, which will be outlined in the following chapters, the highest priority has been given in this manual to the mathematical growth models.

3.3.6 The weight converted von Bertalanffy growth equation

Combining the von Bertalanffy growth equation (Eq. 3.1.0.1)

$$L(t) = L_\infty [1 - \exp(-K(t-t_0))]$$

with the weight/length relationship (Eq. 2.6.1):

$$W(t) = q L(t)^3$$

gives the weight of a fish as a function of age:

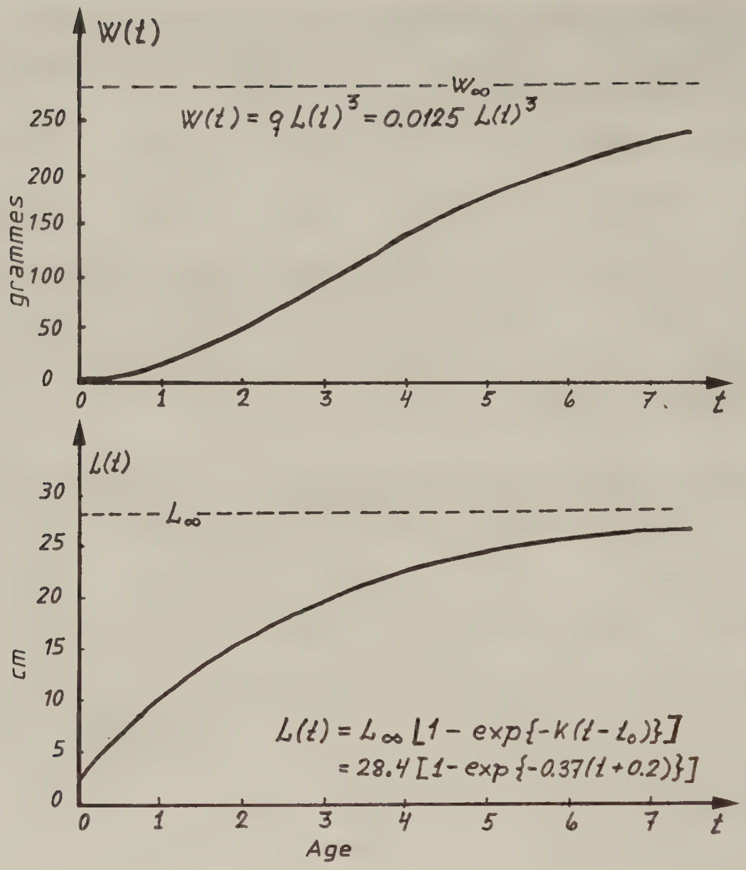


Fig. 3.3.6.1 A length based growth curve converted into a weight based curve

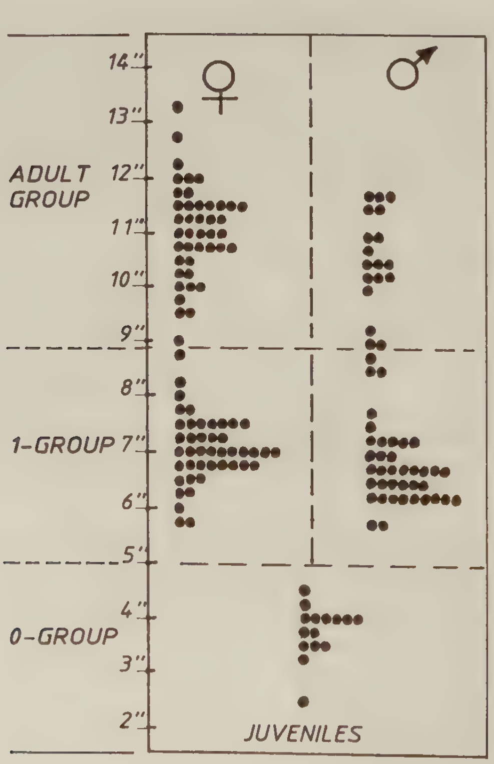


Fig. 3.4.0.1 Length frequency sample of 156 eel pout (Zoarces viviparus) in units of Danish inches. Collected in Holbaek Fjord (Denmark) 10-11 July 1890. (Redrawn from Petersen, 1892)

$$W(t) = q\,L_\infty^3\,[1 - \exp(-K(t-t_o))]^3$$

The "asymptotic weight", W_∞, corresponding to the asymptotic length is (according to Eq. 2.6.1):

$$W_\infty = q\,L_\infty^3$$

The parameter, q, is called the "condition factor". (Note that the letter q is also used in this manual to designate the catchability coefficient, Section 4.3). Thus, "the weight converted von Bertalanffy equation" can be written:

$$W(t) = W_\infty\,[1 - \exp(-K(t-t_o))]^3 \qquad (3.3.6.1)$$

Fig. 3.3.6.1 shows the weight converted growth equation for the von Bertalanffy parameters: $L_\infty = 28.4$ cm, $K = 0.37$ per year, $t_o = -0.2$ years and the condition factor $q = 0.0125$ g per cm^3. These parameters have been estimated for the threadfin bream, *Nemipterus marginatus* in North Borneo waters (Pauly, 1983).

(See **Exercise(s)** in Chapter 17).

3.4 Estimation of age composition from length frequencies

The basic idea behind the techniques to be described in this section dates back to one of the first works on fishery biology, a paper on the eelpout (*Zoarces viviparus*) by Petersen (1892). The length measurements of 156 fish are represented by dots in Fig. 3.4.0.1. Petersen divided the 156 fish into juveniles, males and females and he further divided the adult fish into two size groups:

Medium sized: from 5 to 8 inches
Large sized : from 9 inches and upwards

From earlier observations Petersen knew that in the winter juveniles of about 1.5" could be caught, while in summer the juveniles would all be from 3" to 5" long. Because certain length groups, depending on the season, appeared to be absent, he concluded that the three length groups in the July sample should be interpreted as follows:

less than 5" : 0 -group, born in winter 1889/90
from 5" to 8" : 1 -group, born in winter 1888/89
from 9" and upwards: 2+ -group, born in winter 1887/88 or earlier

(The symbol "a+" stands for the "a-group plus older groups". We call it the "a plus group").

Petersen's findings indicated that *Zoarces viviparus* gives birth once per year during a restricted period. For most temperate species propagation takes place during 2-4 months in winter or spring.

Such a breeding pattern makes it relatively easy to define a cohort. In temperate waters a cohort is simply a yearclass of fish. Because all fish grow at approximately the same rate, a cohort can be followed during the first part of its life by tracing the peaks in the length frequency samples. But when they approach their maximum size this is no longer possible, because by then fish of different ages have reached almost the same length.

Fig. 3.4.0.2 shows a length frequency sample of a tropical species, the coral trout, *Plectropomus leopardus* obtained by Goeden (1978), from Heron Island, Australia. This example appears easy to handle. There are four distinct peaks (A,B,C and D) and it is tempting to interpret these as age

Table 3.4.0.1 Length at age for alternative choices of growth parameters. Plots with observed frequencies are shown in Fig. 3.4.0.2 (column a) and Fig. 3.4.0.3 (columns b and c).

		a *)	b	c	d
	L_∞	57.0	59.5	59.5	70.0
age	K	0.40	0.40	0.34	0.21
	t_o	-0.50	-0.50	-0.65	-1.15
0		10.3	10.8	11.8	15.1
1		25.7	26.8	25.5 ------	25.5
2		36.0	37.6	35.3	33.9
3		42.9	44.8	42.3	40.8
4		47.5	49.7	47.3	46.3
5		50.7	52.9	50.8 ------	50.8
6		52.8	55.1	53.3	54.5
7		54.2	56.5	55.1	57.4

*) see Table 3.1.0.1

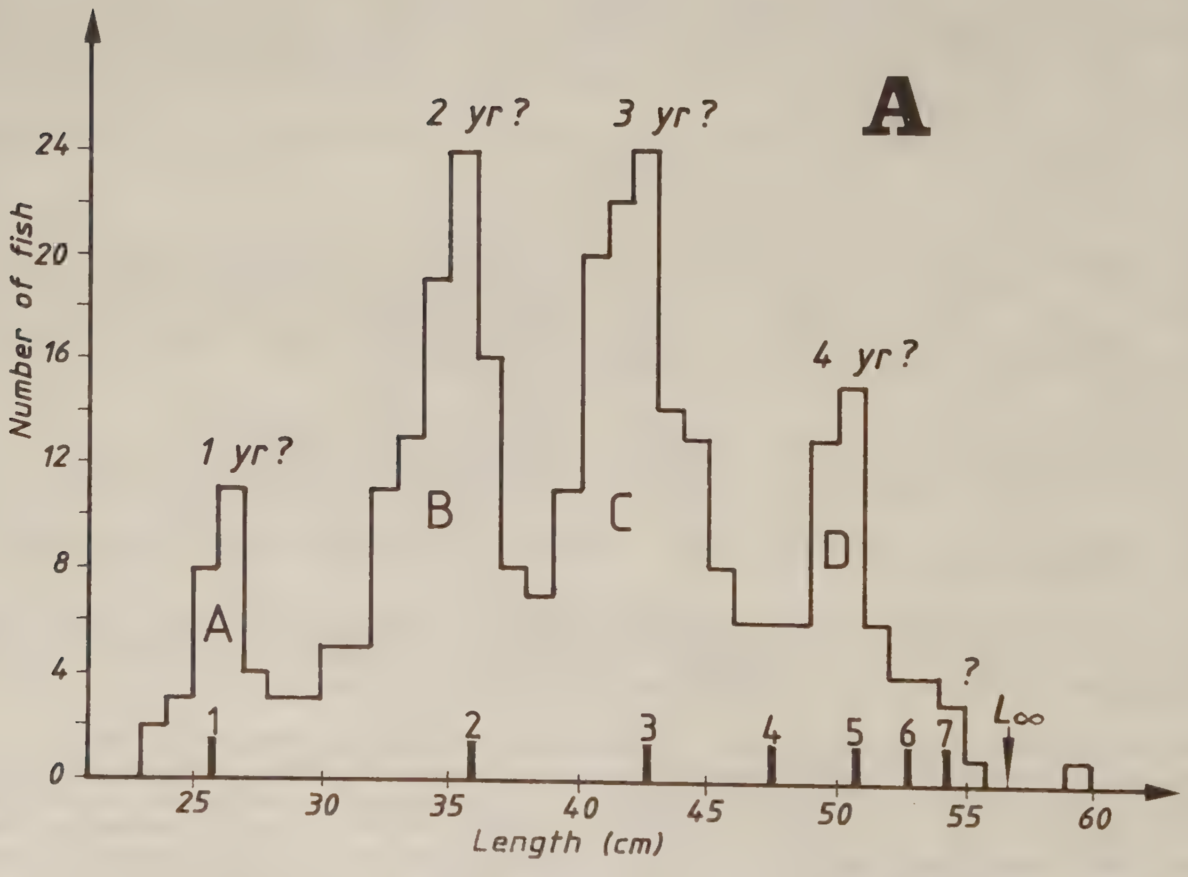


Fig. 3.4.0.2 Length frequency sample of the coral trout (Plectropomus leopardus) from Goeden (1978). The small bars on the length axis indicate the length at age corresponding to the growth parameters: L_∞ = 57 cm, K = 0.4 per year and t_o = -0.5 years (column a of Table 3.4.0.1)

groups 2,3,4 and 5 or, as indicated in Fig. 3.4.0.2, age groups 1,2,3 and 4, as done by Goeden. However, a closer examination shows that none of these interpretations conforms to the von Bertalanffy model. The mean lengths of the peaks B,C and D are approximately 35 cm, 42 cm and 50 cm respectively. When we interpret these as age groups the growth rates become:

Between B and C: $(42-35)/1 = 7$ cm/year
Between C and D: $(59-42)/1 = 8$ cm/year

This does not conform to the von Bertalanffy growth curve, since the growth rate between C and D is expected to be smaller than that between B and C. Thus to give an interpretation conforming to the von Bertalanffy model, peak D must be assigned an age two years older than peak C, and an additional age group must be assumed to exist between peaks C and D. A likely explanation is that peaks C and D represent strong year classes (large number of cohort members), whereas the cohort represented by length groups between peaks C and D stems from a poor year class.

The small solid bars shown on the length axis of Fig. 3.4.0.3 are the lengths at age 1,2,...,7 corresponding to the von Bertalanffy growth curve with the parameters:

$L_\infty = 57$ cm, $K = 0.4$ per year and $t_o = -0.5$ years

The lengths corresponding to the bars were given in Table 3.1.0.2. These parameters interpret the peaks A,B,C and D as the 1-, 2-, 3- and 5-group respectively and place the 4-group between peaks C and D. This particular choice of growth parameters is not based on any fitting techniques or any other rational method. They were derived by a short series of trials with different parameters until a curve was obtained which placed the cohort mean lengths (more or less) where the peaks are, except for age group 4.

Obviously, this is not the only set of growth parameters which produces a certain correspondence to the peaks of Fig. 3.4.0.2. One could have used the greatest length, 59.5 cm, as the estimate of L_∞, for example. Note that considering the definition of L_∞ as the average length of a very old fish, it would not in general be correct to use the largest fish observed as an estimate of L_∞. However, in this case with only 312 fish in the sample the largest fish may give a reasonable estimate of L_∞. Using this L_∞ together with the K value 0.4 and the t_o value -0.5 gives the lengths at age shown in column b of Table 3.4.0.1. Fig. 3.4.0.3B shows the corresponding mean lengths at age together with the length frequency sample. Obviously, this choice of growth parameters produces a less convincing fit to the peaks than the one shown in Fig. 3.4.0.2.

Reducing K to 0.34 and t_o to -0.65 gives a much better agreement between peaks and mean lengths as shown in Fig. 3.4.0.3C. The corresponding mean lengths at age are given in column c of Table 3.4.0.1. Whether this fit is better than the fit shown in Fig. 3.4.0.2 is difficult to assess by visual inspection only.

In general, it is difficult to define a unique solution to this kind of problem. Different values (of L_∞,K,t_o) may produce very similar growth curves. This becomes obvious when one observes that for a given value of L_∞ one can always determine K and t_o so that the curve passes through two prespecified points in the age/length coordinate system.

As an example let us give L_∞ the value 70 cm and let us determine K and t_o so that the curve thus obtained comes close to the curve given in column c of Table 3.4.0.1. We do that by selecting K and t_o so that $L(1) = 25.5$ and $L(5) = 50.8$ cm. From Eq. 3.3.3.3 it follows that:

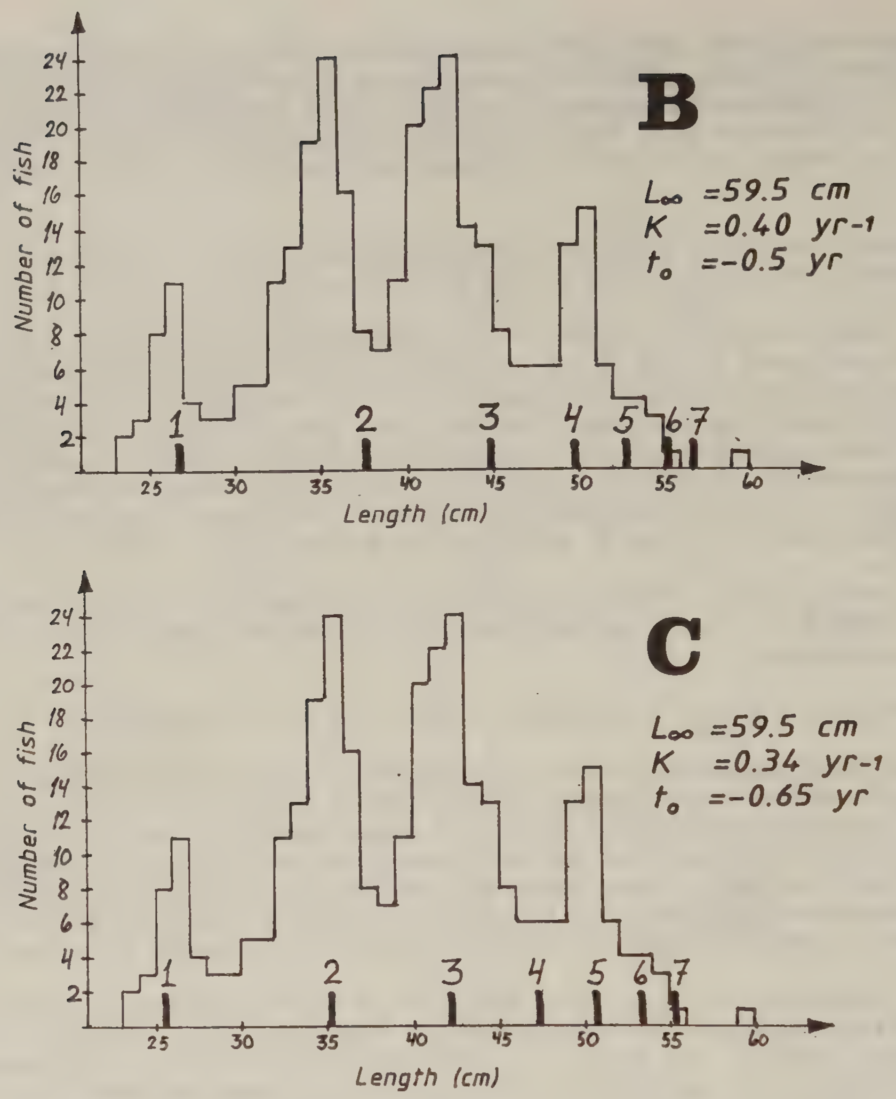


Fig. 3.4.0.3 Length frequency sample of the coral trout (the same data as those of Fig. 3.4.0.2) with length/age superimposed, corresponding to columns b and c of Table 3.4.0.1

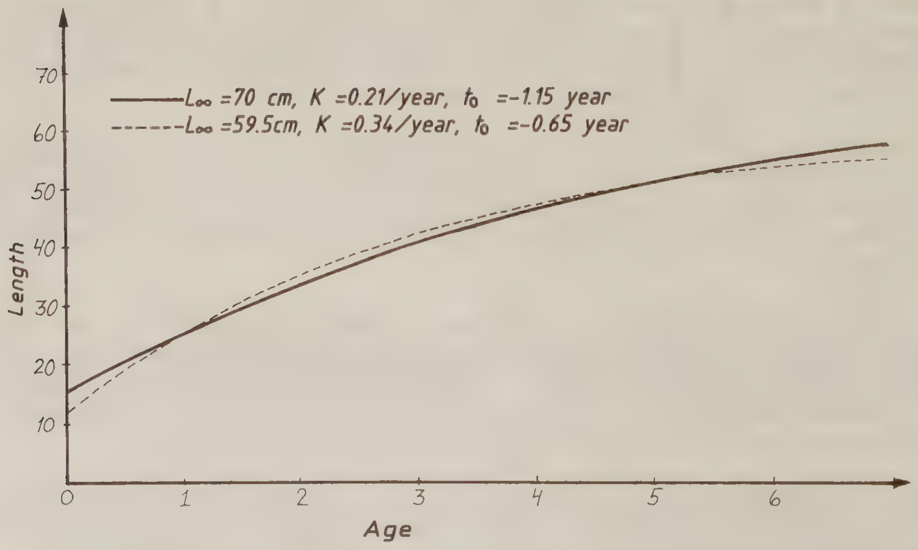


Fig. 3.4.0.4 Example of two growth curves which are approximately equal, but have quite different growth parameters. Derived from columns c and d of Table 3.4.0.1

$$K = \frac{1}{t2-t1} \ln \frac{L\infty - L(t1)}{L\infty - L(t2)} \qquad (3.4.0.1)$$

$$t_o = t1 + \frac{1}{K} \ln(1 - \frac{L(t1)}{L_\infty}) \qquad (3.4.0.2)$$

Thus, for t1 = 1 and t2 = 5 we get:

$$K = \frac{1}{5 - 1} \ln \frac{70.0 - 25.5}{70.0 - 50.8} = 0.21 \text{ per year} \quad \text{and}$$

$$t_o = 1 + \frac{1}{0.21} \ln (1 - \frac{25.5}{70.0}) = -1.15$$

These growth parameters produced the lengths at age shown in column d of Table 3.4.0.1. The two curves are shown in Fig. 3.4.0.4. It appears, that it is difficult to decide which of the two curves gives the best fit to the length frequency sample in Fig. 3.4.0.2.

Fig. 3.4.0.5 shows K and t_o as a function of L_∞ for growth curves fulfilling the condition L(1) = 25.5 cm and L(5) = 50.8 cm. Note that K and t_o decrease as L_∞ increases. Thus, when comparing different estimates of K, L_∞ and t_o the comparison should not be made on the individual parameter but on the resulting growth curves. In the example of columns c and d in Table 3.4.0.1 we would say that the two parameter sets:

$$(L_\infty, K, t_o) = (59.5 , 0.34 , - 0.65) \text{ and}$$

$$(L_\infty, K, t_o) = (70.0 , 0.21 , - 1.15)$$

are approximately equal in the sense that they produce nearly the same growth curves within the range of ages covered. The more old fish in the sample, the better is the estimate of L_∞ and the estimate of K becomes less dependent of the estimate of L_∞.

The discussion above leads up to a warning: do not always consider estimates of growth parameters as related to the physiology of the fish. Only when the sample is large and unbiased you can expect the estimated parameters to reflect their physiological interpretation.

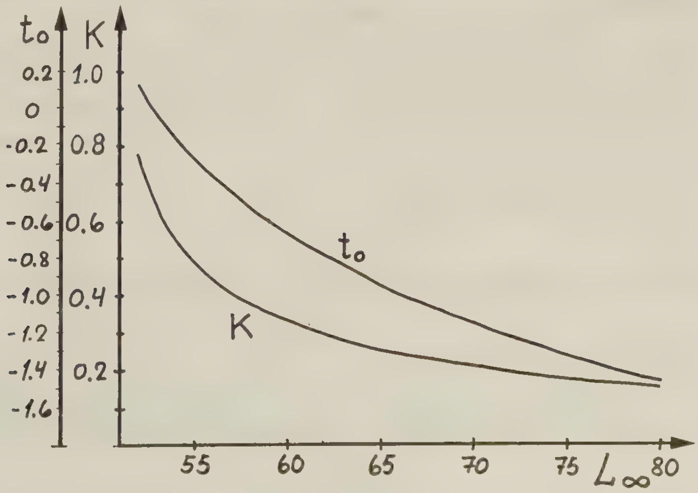


Fig. 3.4.0.5 The growth parameters K and t_o as a function of L_∞ for growth curves fulfilling the condition L(1) = 25.5 cm and L(5) = 50.8 cm

3.4.1 Bhattacharya's method

In Section 2.2 several means of graphical representation of a normal distribution were introduced. One of these is the Bhattacharya (1967) method, which is useful for splitting a composite distribution into separate normal distributions, i.e. when several age groups (cohorts) of fish are contained in the same sample. This method will be discussed in detail based on the hypothetical example of Table 3.2.1.1. In this case we know the solution: the set of normal distributions of which the total is composed. It is therefore possible to check the validity of the result of the analysis.

3.4.1.1 Basis of the computation procedure of the Bhattacharya method

The Bhattacharya method consists basically of separating normal distributions, each representing a cohort of fish, from the overall distribution, starting on the left-hand side of the total distribution. Once the first normal distribution has been determined it is removed from the total distribution and the same procedure is repeated as long as it is possible to

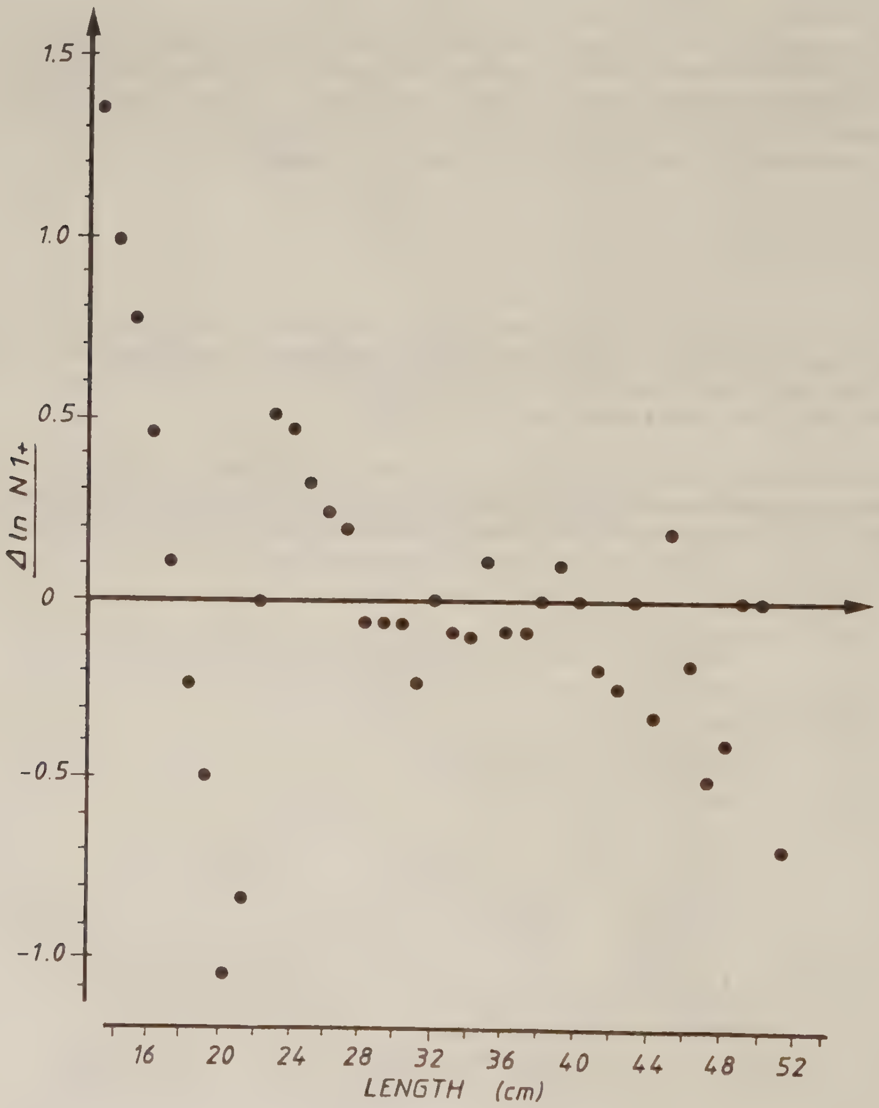


Fig. 3.4.1.1 Bhattacharya method: plot corresponding to columns D and E of Table 3.4.1.1

separate normal distributions from the total distribution. The whole process can be divided into the following steps:

Step 1: Determine an uncontaminated (clean) slope of a normal distribution on the left side of the total distribution.

Step 2: Determine the normal distribution of the first cohort by means of a transformation into a straight line.

Step 3: Determine the numbers of fish per length group belonging to that first cohort and then subtract them from the total distribution.

Step 4: Repeat the process for the next normal distribution from the left, until no more clean normal distributions can be found.

Step 5: Relate the mean lengths of the cohorts determined in steps 1 and 4 to the age difference between the cohorts.

As has already been shown in Section 2.6, a normal distribution is transformed into a straight line when: 1) numbers are replaced by their logarithms and 2) differences are calculated between consecutive logarithmic values. Let N designate the number in a length frequency sample belonging to the length group:

$$[x - dL/2 \; , \; x + dL/2]$$

Where dL is the interval size, x is the interval midpoint and x+dL/2 is the upper limit of the interval.

If a certain length range in the sample contains only one cohort, this part of the frequency sample should conform to a normal distribution (e.g. from 10 cm to 22 cm in the sample shown in Fig. 3.2.2.2). In that case the linear relationship (cf. Eq. 2.6.5):

$$\underline{\Delta} \ln N = a + b\,(x + dL/2)$$

would hold between the difference of logarithms:

$$\underline{\Delta} \ln N = \ln N(x + dL/2 \; , \; x + 3dL/2) - \ln N(x - dL/2 \; , \; x + dL/2)$$

(dependent variable)

and the upper limit of the smallest length group:

$$x + dL/2$$

(independent variable)

(cf. Figs. 2.6.4 and 2.6.5).

Recall that the standard deviation of the normal distribution and the mean are obtained by:

$$s = \sqrt{-dL/b} \quad \text{and} \quad \bar{x} = -\,a/b \quad \text{(cf. Eqs. 2.6.6 and 2.6.7)}$$

3.4.1.2 The computation procedure of the Bhattacharya method

To illustrate the Bhattacharya method we use the example given in Fig. 3.2.2.1 which illustrates the input to the Bhattacharya method and Fig. 3.2.2.2, which shows the normally distributed components. Table 3.4.1.1 columns A to E, shows the calculations required for the Bhattacharya plot.

Column A, labelled "L1-L2", gives the interval limits for the length groups and column B contains the corresponding frequencies. (These frequencies are the same as those given in the "total"-column of Table 3.2.1.1).

Table 3.4.1.1 **Bhattacharya method: estimation of first cohort, N1 (the 1983 spring cohort), and the total minus N1: N2+. "s" in columns A, B and C indicates where to start the calculations of N1 (compare Fig. 3.4.1.2)**

A	B	C	D	E	F	G	H	I
L1-L2	N1+	ln N1+	Δln N1+	L	Δln N1c	ln N1	N1	N2+
12-13	1	0	-	-	-	-	1	0
13-14	4	1.386	1.386	13 *	1.412	-	4	0
14-15	11	2.398	1.012	14 *	1.078	-	11	0
15-16s	24	3.178	0.780	15 *	0.744	("clean")	24	0
16-17	38	3.638s	0.460	16 *	0.410	3.638	38	0
17-18	42s	3.738	0.100	17 *	0.111	3.749	(42.50)	(-0.50)
18-19	33	3.497	-0.241	18 *	-0.205	3.545	(34.64)	(-1.64)
19-20	20	2.996	-0.501	19 *	-0.521	3.024	(20.58)	(-0.58)
20-21	7	1.946	-1.050	20 *	-0.837	2.188	(8.91)	(-1.91)
21-22	3	1.099	-0.847	21	-1.153	1.035	2.82	0.18
22-23	3	1.099	0.000	22	-1.469	-0.434	0.65	2.35
23-24	5	1.609	0.511	23	-1.785	-2.218	0.11	4.89
24-25	8	2.079	0.470	24	-	-	-	8
25-26	11	2.398	0.318	25	-	-	-	11
26-27	14	2.639	0.241	26	-	-	-	14
27-28	17	2.833	0.194	27	-	-	-	17
28-29	16	2.773	-0.061	28	-	-	-	16
29-30	15	2.708	-0.065	29	-	-	-	15
30-31	14	2.639	-0.069	30	-	-	-	14
31-32	11	2.398	-0.241	31	-	-	-	11
32-33	11	2.398	0.000	32	-	-	-	11
33-34	10	2.303	-0.095	33	-	-	-	10
34-35	9	2.197	-0.105	34	-	-	-	9
35-36	10	2.303	0.105	35	-	-	-	10
36-37	11	2.398	0.095	36	-	-	-	11
37-38	10	2.303	-0.095	37	-	-	-	10
38-39	10	2.303	0.000	38	-	-	-	10
39-40	11	2.398	0.095	39	-	-	-	11
40-41	11	2.398	0.000	40	-	-	-	11
41-42	9	2.197	-0.201	41	-	-	-	9
42-43	7	1.946	-0.251	42	-	-	-	7
43-44	7	1.946	0.000	43	-	-	-	7
44-45	5	1.609	-0.336	44	-	-	-	5
45-46	6	1.792	0.182	45	-	-	-	6
46-47	5	1.609	-0.182	46	-	-	-	5
47-48	3	1.099	-0.511	47	-	-	-	3
48-49	2	0.693	-0.405	48	-	-	-	2
49-50	2	0.693	0.000	49	-	-	-	2
50-51	2	0.693	0.000	50	-	-	-	2
51-52	1	0.000	-0.693	51	-	-	-	1

Class interval, dL = 1. Total number in cohort N1: 183.6

*) points used in the regression analysis :

a = 5.4834, b = -0.3160, $\bar{L}1 = -a/b = 17.35$, s1 = $\sqrt{-dL/b}$ = 1.78

Column B is labeled "N1+", which stands for the frequency of all cohorts combined. In general the symbol "Na+" stands for the a'th cohort plus older cohorts. In the present example the cohorts in question are (cf. Table 3.2.1.1):

Cohort no. 1: Spring cohort recruited in 1983
Cohort no. 2: Autumn cohort recruited in 1982
Cohort no. 3: Spring cohort recruited in 1982
Cohort no. 4: Autumn cohort recruited in 1981
...... etc.

"N3+" stand for the cohorts recruited in the spring of 1982 and earlier. Column C gives the logarithms of N1+. Column D gives Δ ln N1+ or, the ln N1+ of the current line minus the ln N1+ of the foregoing line.

As an example consider the length groups 14-15 cm and 15-16 cm:

A	B	C	D	E
L1-L2	N1+	ln N1+	Δ ln N1+	L
14-15	11	2.398 = ln 11		
15-16	24	3.178 = ln 24	0.780 = ln 24 - ln 11	15

Recall from Section 2.6 that Δ ln N1+ should be plotted against the upper limit of the smallest length group of the two from which Δ ln N1+ is calculated. This is the length given in column E.

The Bhattacharya plot corresponding to columns D and E is given in Fig. 3.4.1.1. In this particular case we know the solution to the problem as the data were calculated from a set of parameters. The solution is given in the three last lines of Table 3.2.1.1. Thus the first component (the 1983 spring cohort) should have a mean value of 17.3 cm and a standard deviation of 1.7 cm. Visual inspection of Fig. 3.4.1.1 shows that a straight line can be fitted to the first eight points (indicated by "*" in Table 3.4.1.1). This straight line corresponds to the normally distributed component, which we interpret as the N1 cohort (the 1983 cohort). That the straight line corresponding to N1 comes out so nicely is not surprising, considering the length frequency diagrams of Figs. 3.2.2.1 and 3.2.2.2, where the first component has very little overlap with its neighbouring component.

Carrying out the regression analysis of Δ ln N1+ (column D) on length (column E) yields the results:

a (intercept) = 5.4834   b (slope) = - 0.3160

from which the mean length, $\bar{L}1$, and the standard deviation of cohort N1 are calculated (cf. Section 2.6):

$\bar{L}1 = -a/b = 17.35$ $s = \sqrt{-1/b} = 1.78$

The regression line is shown in Fig. 3.4.1.2.

The next step is to estimate the number of fish by length group belonging to cohort N1. From each length group of the total sample we then subtract the part belonging to the N1 cohort to get N2+. It then should become easier to identify the next cohort N2 and separate it from N2+, since it is no longer contaminated by N1. Subsequently the whole procedure described above is repeated to separate N2 from N2+. The procedure is shown in columns F to H in Table 3.4.1.1. We start by selecting a "clean" length group, i.e. a length group we believe consists exclusively of observations from cohort N1. In this case the 38 observations from length group 16-17 cm have been selected. We could have chosen the first length group, 12-13 cm, arguing that this is more likely not to be contaminated by N2. However, because it contains only one observation it is expected to be more uncertain (to have a larger relative variance) than a length group with many observations.

In columns B and C we found that ln N1(16-17) = ln 38 = 3.638. The number of observations in the next length group, 17-18 cm, was 42. The "theoretical" or "calculated" number, i.e. the number of fish in this length group which corresponds exactly to the estimated regression line (Fig. 3.4.1.2) is slightly different. By calculating the point on the regression line corresponding to length 17 cm we get:

/\ ln N1c(17) = a + b * 17 = 5.4834 - 0.3160 * 17 = 0.111

(where the suffix "c" to N1 stands for "calculated value").

The calculated ln(number) in length group 17-18 cm then becomes:

ln N1c(17-18) = ln N1(16-17) + /\ ln N1c(17)

= 3.638 + 0.111 = 3.749

(see column G in Table 3.4.1.1). The corresponding calculated number is:

N1c(17-18) = exp(3.749) = 42.5 (see column H in Table 3.4.1.1)

For the 18-19 cm group we find in the same manner:

N1c(18) = 5.4834 - 0.3160 * 18 = -0.205

ln N1c(18-19) = ln N1c(17-18) + /\ ln N1c(18)

= 3.749 - 0.205 = 3.544

N1c(18-19) = exp(3.544) = 34.64 (see column H)

The results are not exactly the same as the observed frequencies given in column B, because observations always deviate somewhat from the theoretical values. In the present case of a hypothetical data set, the deviations are due to rounding errors. With "real" data there are also deviations caused by "random noise". Even if the sample is a perfect random sample the observations will fluctuate around the true length distribution (the length distribution of the population).

A continuation of this process produces columns G and H in Table 3.4.1.1. However, the downward calculation process in columns G and H is stopped when the number in cohort N1 becomes negligible. For length group 23-24 cm the number in the N1 cohort is only 0.11, which is considered to be approximately equal to zero. The figures at which the calculation process is started are indicated in the table by "s".

Finally, the figures in column I of Table 3.4.1.1 are obtained by subtracting N1 from N1+, yielding N2+:

N2+ = N1+ - N1 or column I = column B - column H

As N2+ is cohort no. 2 plus older cohorts it makes no sense to assign a negative number to column I as is the case for the figures in brackets in Table 3.4.1.1. In practice it may well happen that the figures in column I become negative because of random varation of the observations. However, this can be adjusted. Whenever the estimate of N2+ numbers becomes negative we assign the value zero to N2+ (in column I) while the N1 is given the value of column B. As a result of this adjustment the first rows in Table 3.4.1.1 will actually read:

A	B	C	D	E	F	G	H	I
L1-L2	N1+	ln N1+	/\ln N1+	L	/\ln N1c	ln N1	N1	N2+
	....	....						
16-17	38	3.638s	0.460	16 *	0.410	3.638	38	0
17-18	42	3.738	0.100	17 *	0.111	3.749	42	0
18-19	33	3.497	-0.241	18 *	-0.205	3.545	33	0
19-20	20	2.996	-0.501	19 *	-0.521	3.024	20	0
20-21	7	1.946	-1.050	20 *	-0.837	2.188	7	0
21-22	3	1.099	-0.847	21	-1.153	1.035	2.82	0.18
	....	....						

Hereby we come to the end of the use of Table 3.4.1.1. By eliminating all values pertaining to N1 we create the new Table 3.4.1.2, with N2+ as column B. Column I of Table 3.4.1.1 has been transferred to column B of Table 3.4.1.2. The whole procedure can then be repeated.

Fig. 3.4.1.3 shows the Bhattacharya plot for N2+ together with the estimated line for N1. Only the points to the right of the broken line of Fig. 3.4.1.2 are used in the analysis now. The N1-line is shown for comparison only. Some points have been moved due to the subtraction of N1. The "old" points (i.e. those from the N1+ plot) are indicated by "X" and the "new" points by a dot, in cases where the movement is visible. The two first points (corresponding to lengths 21 and 22 cm) are disregarded, because they refer to very small numbers of individuals.

The selection of points to fit a straight line is now a bit more difficult than in the case of the first cohort. In Fig. 3.4.1.3 the six points from lengths 23 cm to lengths 28 cm were chosen. One can question why these points were chosen in preference to the points, for example, from length 24 to 29 cm or from length 24 to 28 cm? The choice is a subjective one. The results of the Bhattacharya method may sometimes be dependent on the person who actually performs the analysis. If, for example, only the points from lengths 24 to 27 cm were used the estimated mean length would be 28.7 cm and the standard deviation 3.2 cm. The actual choice made in Fig. 3.4.1.3 gives a mean value of L2 = 27.76 cm and a standard deviation of 2.66 cm which are both very close to the true values (see Tables 3.2.1 and 3.4.1.2). However, this cannot be used as a justification for this choice, because in real life we would not know what the true values should be. Also the selection of the "clean" value of ln N2 from which N2 and N3+ are calculated is a subjective one. The more the observations deviate from the calculated frequencies the more pronounced the element of subjectivity.

In summary, the results obtained so far are:

cohort N1:	mean length 17.35 cm, standard deviation 1.78 cm (Table 3.4.1.1)
cohort N2:	mean length 27.76 cm, standard deviation 2.66 cm (Table 3.4.1.2)

Now that the two first mean lengths of cohorts have been estimated we are in a position to obtain a first rough estimate of the von Bertalanffy parameter K, provided we also have an estimate of the age difference between the two cohorts. We use Eqs. 3.4.0.1 and 3.4.0.2 with a time difference between the two cohorts equal to t2-t1 = 0.5 year. Further a rough estimate of L_∞ is obtained from the length frequency sample, which tells

Table 3.4.1.2 Bhattacharya method: Estimation of the second cohort, N2 (the 1982 autumn cohort), and the total minus N2, N3+ (compare Fig. 3.4.1.3)

A	B	C	D	E	F	G	H	I
L1-L2	N2+	ln N2+	/\ln N2+	L	/\ln N2c	ln N2	N2	N3+
12-13	0	-	-	-	-	-	0	0
13-14	0	-	-	13	-	-	0	0
14-15	0	-	-	14	-	-	0	0
15-16	0	-	-	15	-	-	0	0
16-17	0	-	-	16	-	-	0	0
17-18	0	-	-	17	-	-	0	0
18-19	0	-	-	18	-	-	0	0
19-20	0	-	-	19	-	-	0	0
20-21	0	-	-	20	-	-	0	0
21-22	0.18	(-1.715)	(-0.175)	21	-	-	0.18	0
22-23	2.35	0.845	2.569	22	-	-	2.35	0
23-24	4.89	1.587	0.733	23 *	-	-	4.89	0
24-25	8	2.079	0.492	24 *	-	-	8	0
25-26s	11	2.398	0.318	25 *	-	("clean")	11	0
26-27	14	2.639s	0.241	26 *	0.245	2.639	14	0
27-28	17s	2.833	0.194	27 *	0.107	2.746	15.48	1.41
28-29	16	2.773	-0.061	28 *	-0.034	2.712	15.06	0.94
29-30	15	2.708	-0.065	29	-0.175	2.536	12.63	2.36
30-31	14	2.639	-0.069	30	-0.317	2.220	9.21	4.79
31-32	11	2.398	-0.241	31	-0.458	1.762	5.82	5.18
32-33	11	2.398	0.000	32	-0.599	1.163	3.20	7.80
33-34	10	2.303	-0.095	33	-0.740	0.423	1.53	8.47
34-35	9	2.197	-0.105	34	-0.881	-0.459	0.63	8.36
35-36	10	2.303	0.105	35	-1.023	-1.481	0.23	9.77
36-37	11	2.398	0.095	36	-1.164	-2.645	0.07	10.93
37-38	10	2.303	-0.095	37	-	-	-	10
38-39	10	2.303	0.000	38	-	-	-	10
39-40	11	2.398	0.095	39	-	-	-	11
40-41	11	2.398	0.000	40	-	-	-	11
41-42	9	2.197	-0.201	41	-	-	-	9
42-43	7	1.946	-0.251	42	-	-	-	7
43-44	7	1.946	0.000	43	-	-	-	7
44-45	5	1.609	-0.336	44	-	-	-	5
45-46	6	1.792	0.182	45	-	-	-	6
46-47	5	1.609	-0.182	46	-	-	-	5
47-48	3	1.099	-0.511	47	-	-	-	3
48-49	2	0.693	-0.405	48	-	-	-	2
49-50	2	0.693	0.000	49	-	-	-	2
50-51	2	0.693	0.000	50	-	-	-	2
51-52	1	0.000	-0.693	51	-	-	-	1

Total number in cohort N2: 104.4

*) points used in the regression analysis :

a = 3.9194, b = -0.1412, $\bar{L}2 = -a/b = 27.76$, $s2 = \sqrt{-dL/b} = 2.66$

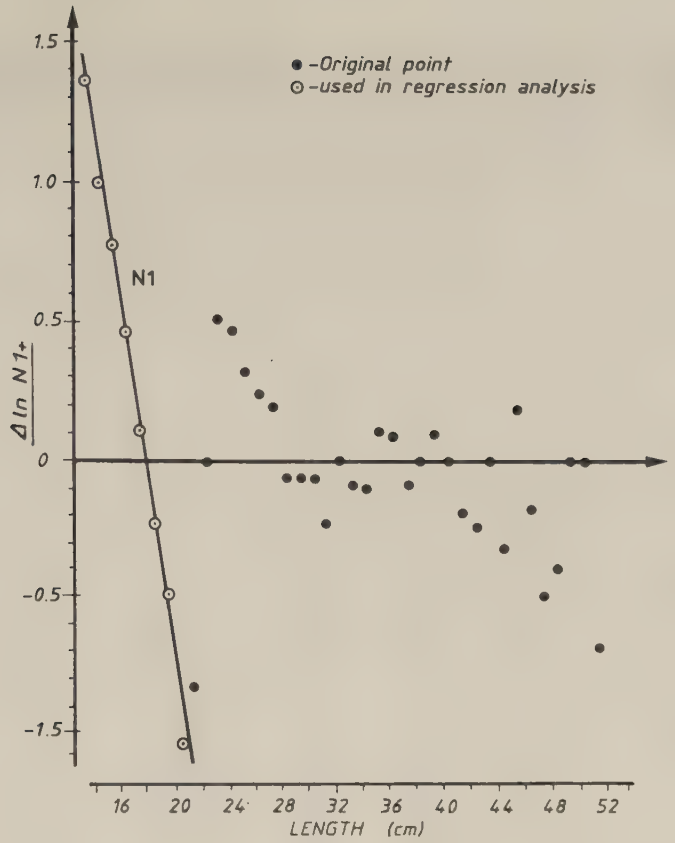


Fig. 3.4.1.2 Bhattacharya method: regression line estimated for the first cohort (compare to columns D and E in Table 3.4.1.1)

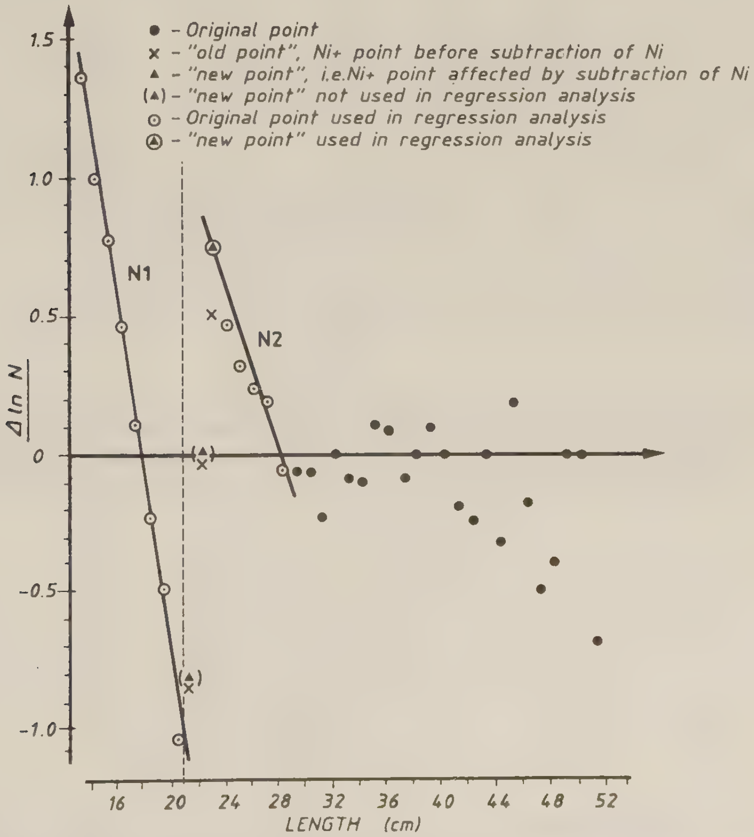


Fig. 3.4.1.3 Bhattacharya method: regression line estimated for the second cohort (compare columns D and E in Table 3.4.1.2)

Table 3.4.1.3 **Bhattacharya method: Estimation of third cohort, N3 (the 1982 spring cohort), and the total minus N3, N4+ (compare Fig. 3.4.1.4)**

A	B	C	D	E	F	G	H	I
L1-L2	N3+	ln N3+	Δln N3+	L	Δln N3c	ln N3	N3	N4+
....								
25-26	0	-	-	25	-	-	0	0
26-27	0	-	-	26	-	-	0	0
27-28	1.41	(0.344)	(0.344)	27	-	-	1.41	0
28-29	0.94	(-.062)	(-.405)	28	-	-	0.94	0
29-30	2.36	(0.859)	(0.921)	29	-	-	2.36	0
30-31	4.79	1.567	0.708	30 *	-	-	4.79	0
31-32s	5.18	1.645	0.078	31 *	-	("clean")	5.18	0
32-33	7.80	2.054s	0.409	32 *	-	2.504	7.80	0
33-34	8.47s	2.137	0.082	33 *	0.109	2.163	8.47	0
34-35	8.36	2.123	-0.013	34 *	-0.035	2.128	8.36	0
35-36	9.77	2.279	0.156	35	-0.179	1.950	7.03	2.74
36-37	10.93	2.392	0.112	36	-0.322	1.627	5.09	5.84
37-38	10	2.303	-0.089	37	-0.466	1.161	3.19	6.81
38-39	10	2.303	0.000	38	-0.610	0.551	1.73	8.26
39-40	11	2.398	0.095	39	-0.754	-0.203	0.82	10.18
40-41	11	2.398	0.000	40	-0.898	-1.100	0.33	10.67
41-42	9	2.197	-0.201	41	-1.041	-2.142	0.12	8.88
42-43	7	1.946	-0.251	42	-	-	-	7
43-44	7	1.946	0.000	43	-	-	-	7
44-45	5	1.609	-0.336	44	-	-	-	5
45-46	6	1.792	0.182	45	-	-	-	6
46-47	5	1.609	-0.182	46	-	-	-	5
47-48	3	1.099	-0.511	47	-	-	-	3
48-49	2	0.693	-0.405	48	-	-	-	2
49-50	2	0.693	0.000	49	-	-	-	2
50-51	2	0.693	0.000	50	-	-	-	2
51-52	1	0.000	-0.693	51	-	-	-	1

Total number in cohort N3: 104.4

*) points used in the regression analysis :

a = 4.8544, b = -0.1438, $\bar{L}3 = -a/b = 33.76$, s3 = $\sqrt{-dL/b}$ = 2.64

Table 3.4.1.4 **Estimation of K and t_o by the von Bertalanffy plot using arbitrary input ages and the mean lengths estimated in Tables 3.4.1.1 to 3.4.1.3 (Compare Table 3.3.3.1)**

x		y
t	$\bar{L}(t)$	$-\ln(1 - \bar{L}(t)/50)$
0.5	17.4	0.428
1.0	27.8	0.812
1.5	33.8	1.127

a (intercept) = 0.09
b (slope) = 0.699 = K
$t_o = - a/b = - 0.13$

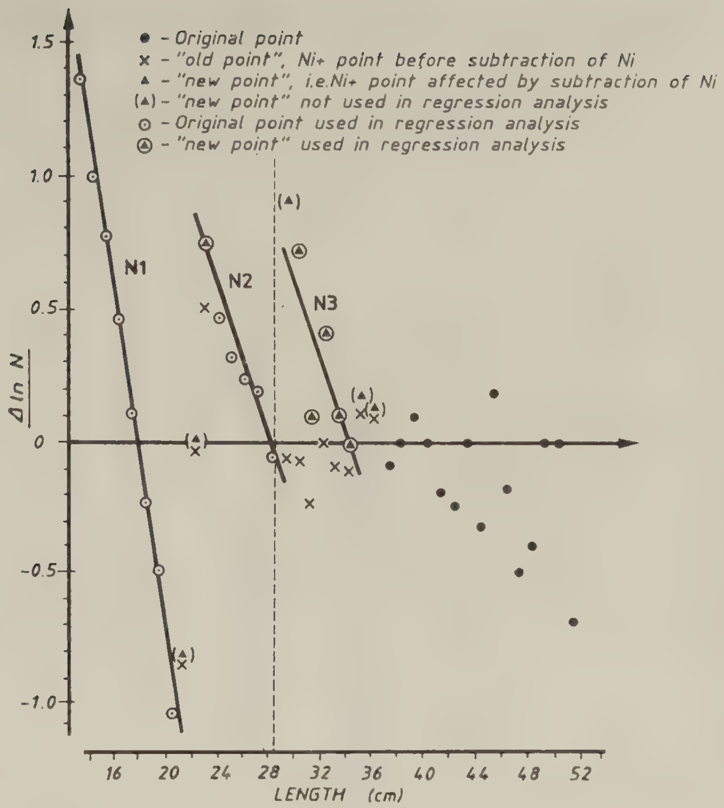


Fig. 3.4.1.4 Bhattacharya method: regression line estimated for the third cohort (compare columns D and E in Table 3.4.1.3)

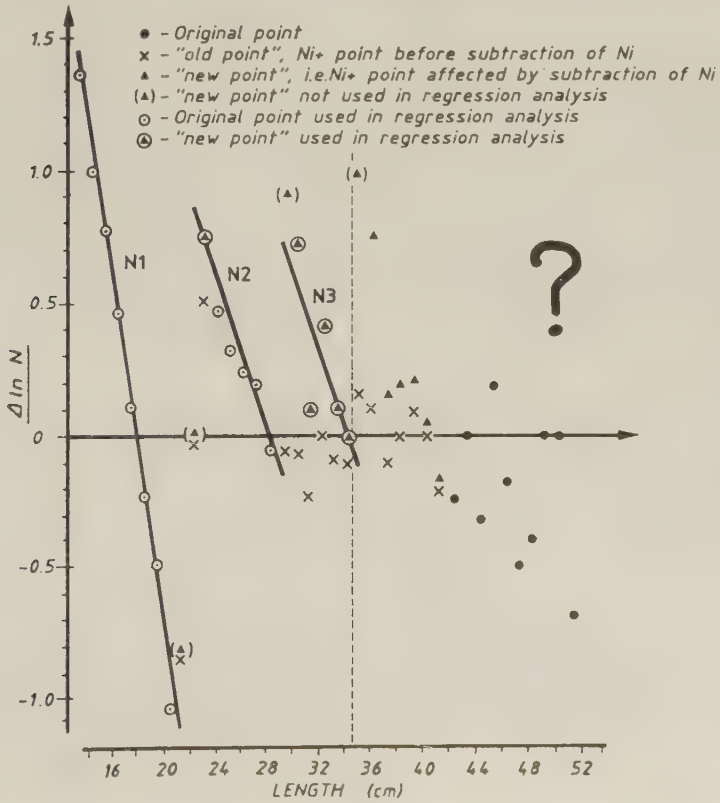


Fig. 3.4.1.5 Bhattacharya method: plot for estimation of the fourth cohort (compare columns D and E in Table 3.4.1.5). In this case the fit was considered too dubious

Table 3.4.1.5 Bhattacharya method: Attempt to estimate the fourth cohort, N4 (the 1981 autumn cohort). Compare Fig. 3.4.1.5

A	B	C	D	E	F	G	H	I
L1-L2	N4+	ln N4+	/\ln N4+	L	/\ln N4c	ln N4	N4	N5+
	....	....	....					
34-35	0	-	-	34	-	-	0	0
35-36	2.74	(1.001)	(1.001)	35	?	?	?	?
36-37	5.84	1.765	0.757	36	?	?	?	?
37-38	6.84	1.918	0.154	37 (*)	?	?	?	?
38-39	8.26	2.111	0.193	38 (*)	?	?	?	?
39-40	10.18	2.320	0.209	39 (*)	?	?	?	?
40-41	10.67	2.367	0.047	40 (*)	?	?	?	?
41-42	8.88	2.184	-0.184	41 (*)	?	?	?	?
42-43	6.96	1.940	-0.244	42	?	?	?	?
43-44	7	1.946	0.006	43	?	?	?	?
44-45	5	1.609	-0.336	44	?	?	?	?
45-46	6	1.792	0.182	45	?	?	?	?
46-47	5	1.609	-0.182	46	?	?	?	?
47-48	3	1.099	-0.511	47	?	?	?	?
48-49	2	0.693	-0.405	48	?	?	?	?
49-50	2	0.693	0.000	49	?	?	?	?
50-51	2	0.693	0.000	50	?	?	?	?
51-52	1	0.000	-0.693	51	?	?	?	?

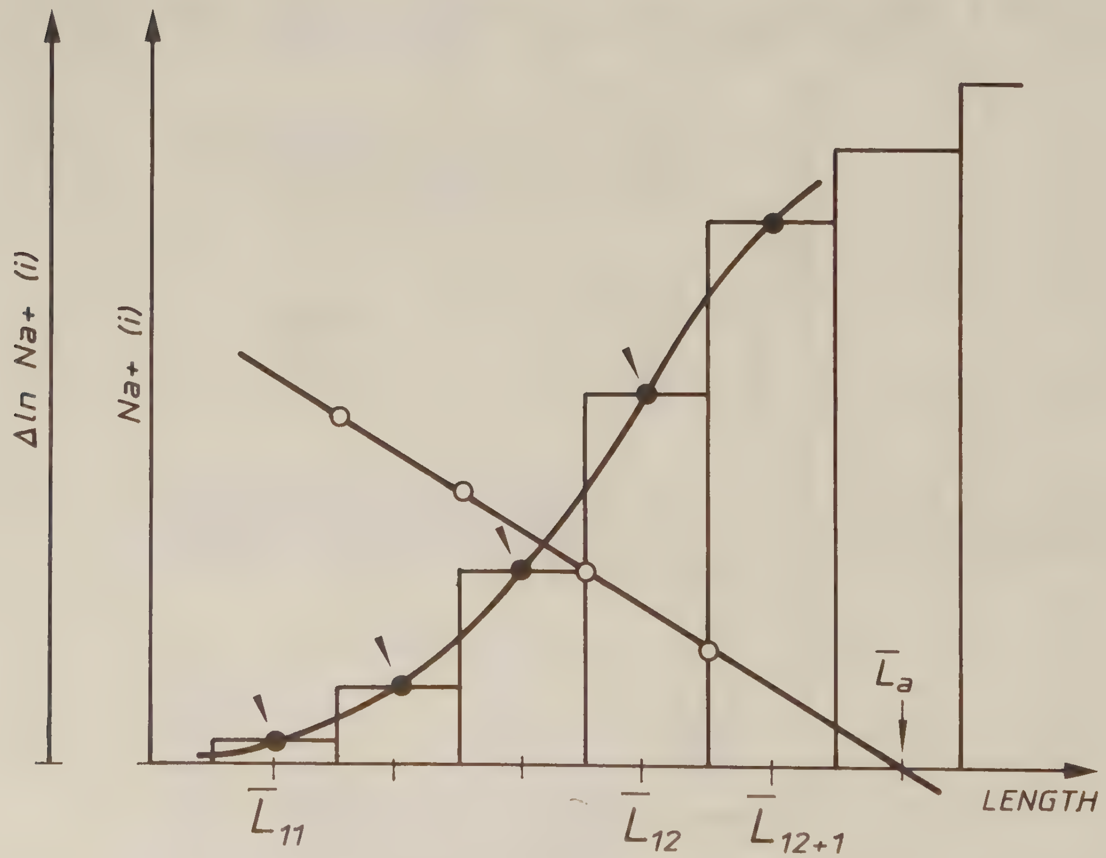


Fig. 3.4.1.6 The points used to estimate cohort numbers by Eq. 3.4.1.1 (for further explanation, see text)

us that the fish rarely get longer than 50 cm, so L_∞ = 50 cm. From Eq. 3.4.0.1 we get:

$$K = \frac{1}{t2-t1} \ln \frac{L\infty - \bar{L}(t1)}{L\infty - \bar{L}(t2)} = \frac{1}{0.5} \ln \frac{50 - 17.4}{50 - 27.8} = 0.77$$

and from Eq. 3.4.0.2:

$$t_o = t1 + \frac{1}{K} \ln(1 - \frac{\bar{L}(t1))}{L_\infty} = 0.5 + \frac{1}{.77} \ln(1 - \frac{17.4}{50}) = -0.06$$

where the value t1 = 0.5 is an arbitrary age.

Thus, as a first rough estimate of the growth curve we have

$$L(t) = 50 [1 - \exp(-0.77(t + 0.06))]$$

The estimation procedure presented above is not generally recommended. It has been given here to demonstrate how little data are actually required to roughly estimate a growth curve. Such a first estimate, however, may be used to predict the next mean length, i.e. the mean length of cohort N3. Assuming cohort N3 to be 1.5 years old we get:

$$L(1.5) = 50 [1 - \exp(-0.77(1.5 + 0.06))] = 35.0 \text{ cm}$$

Table 3.4.1.3 has been prepared for the analysis of N3+ and the related Bhattacharya plot is shown in Fig. 3.4.1.4. The selection of points used for the regression for N3 is now even more questionable than the one made for cohort N2. However, the estimated mean value of $\bar{L}3$ = 33.8 cm came out reasonably well compared to the value of 35 cm calculated above and which we happen to know is true (Table 3.2.1.1).

With three mean lengths we are now in a position to apply the von Bertalanffy plot (cf. Section 3.3.3, Eq. 3.3.3.1), again assuming the arbitrary age of 0.5 years for the first cohort. The input data for the estimation of K and t_o from the von Bertalanffy plot are shown in Table 3.4.1.4, together with the results of the regression analysis. The estimate of t_o is an arbitrary value, since we used arbitrary ages. Nevertheless it puts us in a position to calculate length at age because the shape of the growth curve is independent of t_o. With the new growth parameters estimated in Table 3.4.1.4 the expected mean length of cohort N4 becomes:

$$L(2.0) = 50 [1 - \exp(-0.7 (2.0+0.13))] = 38.7 \text{ cm}$$

We now continue with the Bhattacharya method to estimate N4. Table 3.4.1.5 and Fig. 3.4.1.5 show the Bhattacharya analysis for N4+. It is difficult to see a straight line. Selecting the five points corresponding to lengths 37-41 cm would give a mean length of 40.0 cm, which is a reasonable value. (Actually, this value is very close to the true value of 40.2 cm (cf. Table 3.2.1), but we are not supposed to have that information).

At this stage one should probably consider the fit of a straight line as being so poor that the analysis should be terminated. When to stop is largely a matter of taste, although some objective criteria for limitations of the Bhattacharya method can be devised as will be discussed in Section 3.5.4. Anyway, we stop here to bring this example to and end.

Table 3.4.1.6 **Application of Eq. 3.4.1.1 for estimation of N1 and N2. Figures in brackets are the estimates from Tables 3.4.1.1 and 3.4.1.2**

A: $\bar{L}1 = 17.35$, $s1 = 1.78$

length group	$\bar{L}i$	$\frac{dL}{s1\sqrt{2\pi}}\left[\exp - \frac{(\bar{L}i-\bar{L}1)^2}{2\ s1^2}\right]$	observed n1+	calculated n1
13 - 14	13.5	0.0216	4	4.0
14 - 15	14.5	0.0622	11	11.4
15 - 16	15.5	0.1306	24	24.0
16 - 17	16.5	0.2000	38	36.7
17 - 18	17.5	0.2233	42	41.0
18 - 19	18.5	0.1819	33	33.4
19 - 20	19.5	0.1081	20	19.9
20 - 21	20.5	0.0468	7	8.6
	sum	0.9745 (=X)	179 (=Y)	179.0

N1 = X/Y = 179/0.9745 = 183.7 (183.7)

B: $\bar{L}2 = 27.76$, $s2 = 2.66$

length group	$\bar{L}i$	$\frac{dL}{s2\sqrt{2\pi}}\left[\exp - \frac{(\bar{L}i-\bar{L}2)^2}{2\ s2^2}\right]$	observed n1+	calculated n1
23 - 24	23.5	0.0416	4.9	4.6
24 - 25	24.5	0.0708	8	7.8
25 - 26	25.5	0.1045	11	11.5
26 - 27	26.5	0.1341	14	14.7
27 - 28	27.5	0.1493	17	16.4
28 - 29	28.5	0.1443	16	15.9
	sum	0.6446 (=X)	70.9(=Y)	70.9

N2 = X/Y = 70.9/0.6446 = 110.0 (104.4)

3.4.1.3 More features of the Bhattacharya analysis

1) Improvements in estimating the size of a cohort

One of the weak points in the paper and pencil version of the Bhattacharya method presented above is that of the estimation of the sizes of the components, i.e. the number of fish in a cohort, since this calculation is based on the (subjective) selection of one "clean point", from which the values of ln Na+ are calculated.

A more rigorous statistical approach would be to apply all points used for the estimation of the line. It works as follows: Let component "a" be the first component in the current state of the exercise (i.e. components 1,...,a-2, a-1 have already been removed) and let

$$n_{a+}(i), \quad i = j1,...,j2+1$$

be the "observations" used in the regression (see Fig. 3.4.1.6). Thus the suffix "a+" symbolizes the components of age a and older. If the observations $n(a+)(j1)$, $(j1+1)$, ..., $n_{a+}(j2)$ are clean then we expect their sum to equal the expected number in the section from j1 to j2 of the normal distribution we are dealing with.

$$Y = Na\ X \qquad (3.4.1.1)$$

where:

$$X = \frac{dL}{sa\sqrt{2\pi}} \sum_{i=j1}^{j2} \exp - \frac{(\bar{L}i-\bar{L}a)^2}{2\ sa^2}, \qquad Y = \sum_{i=j1}^{j2} n_{a+}(i) \quad \text{and}$$

Na = total number of fish in component a
$\bar{L}i$ = interval midpoint of interval no. i
$\bar{L}a$ = mean length estimated for component a
sa = standard deviation estimated for component a

As both X and Y in Eq. 3.4.1.1 are known Na is estimated by Y/X. We have thus obtained a way of estimating Na where all "clean" points are used. Note (see Fig. 3.4.1.6) that the last point, j2+1, is not used. The reason is that this point is probably "contaminated" by the next component, while at the same time it is carrying most weight in the regression for the estimation of Na. In order to be on the safe side we exclude it.

Tables 3.4.1.6A and B show the procedure for estimation of N1 and N2. It appears that in this case the estimates of N1 and N2 are only slightly different from those obtained by selecting a "clean starting point". However, the main advantage of the procedure where all points (except for the last one) are used, is that it does not require the subjective choice of a "clean starting point". The disadvantage is that it requires more calculations.

2) Bias

Input data for the Bhattacharya analysis are often biased due to gear selection and recruitment, i.e. the small fish are underrepresented in the frequency samples, either because they escape through the meshes of the gear, or because they have not yet migrated from the nursery gorunds to the fishing grounds (cf. Section 2.7). Aspects connected with bias caused by selection will be discussed in Chapter 6, where also a method to adjust length frequency samples for selection will be presented. In many cases the Bhattacharya analysis should be preceded by an adjustment for selection.

Another source of bias is observed for migratory fish species. Sometimes components are lacking because the cohort was not present in the area where the samples were taken. This aspect will be discussed in Chapter 11.

3) Computer programs

As you may have noticed, the Bhattacharya exercise takes some time to do by "paper and pencil". With the aid of a computer (which may be a microcomputer), however, the method is not hard to work with in practice.

The computer program "BHATTAC" in the LFSA package of microcomputer programs (see Chapter 14) closely follows the set-up explained in the foregoing. With a little experience you can do the exercise of Section 3.4.1 with BHATTAC in a few minutes. The program has a number of additional features: Whenever you have estimated a component BHATTAC displays a graph like Fig. 3.2.2.2 on the screen, allowing you to evaluate the fit to the original data. BHATTAC also checks whether your results are reasonable or not by calculating the "separation index", described in Section 3.5.4. Perhaps the most important feature of BHATTAC, compared to the "paper-and-pencil-method", is that it allows you to do the analysis several times, each time with a different set of input data. You may for example want to try out a range of alternative ways to fit the straight lines in the Bhattacharya plot.

When doing the Bhattacharya analysis on the computer you should always, as a matter of routine, try out different length class sizes (cf. Exercise 3.4.1), since it often happens that the structure of the points on the Bhattacharya plot emerges only for an optimal length class size, which you may find simply by trying out various alternatives. Similar improvements may be obtained by pooling data over longer periods. In most cases you will be working with time series of length frequencies (to be dealt with in Section 3.4.2), for example in the form of monthly length frequencies. You will then have the choice between, say, working with samples representing one month or to pool the data of three months to represent a quarter of the year. Such alternative aggregations of the basic data can easily be made by computer.

The "COMPLEAT ELEFAN" package of microcomputer programs (Gayanilo, Soriano and Pauly, 1988) contains a program "MPA", a part of which also does the Bhattacharya exercise. "MPA" and "BHATTAC" work essentially in the same way.

Pauly and Caddy (1985), have developed a slightly different version of the Bhattacharya method for use with the HP programmable calculator. In their version the lines are determined by three successive points only, which are chosen so that they have the highest negative correlation coefficient. Their version is an attempt to turn the Bhattacharya method into an objective method, i.e. a method producing results independent of the person carrying out the analysis.

(See **Exercise(s)** in Chapter 17).

3.4.2 Modal progression analysis

The example used in the preceding section was based on one length freqency sample collected during one survey. It was demonstrated that a somewhat rough estimate of the growth equation could be obtained from such a data set. The pair (L_∞, K) could be estimated, whereas t_o could not.

Now suppose we had the type of data described in Section 3.2.2, i.e. length frequency samples from each month or quarter during one or several years. The example illustrated in Fig. 3.2.1.1 consists of 12 length frequency samples collected in January, April, July and October during three years (1982 to 1984). Such a time series puts us in a much better position to estimate growth parameters than in the case of a single sample (October

Table 3.4.2.1 **Results of Bhattacharya analyses of the time series of length frequency samples illustrated in Fig. 3.2.1.1**

date of sample	third component	second component	first component
JAN 82	27.9	23.5	9.8
APR 82	32.0	28.1	16.5
JUL 82	31.8	23.1	8.0
OCT 82	34.6	28.0	15.3
JAN 83	32.0	21.8	10.0
APR 83	35.1	27.0	16.5
JUL 83	30.9	23.5	9.2
OCT 83 *)	33.8	27.8	17.4
JAN 84	32.9	24.0	8.3
APR 84	-	28.2	16.8
JUL 84	-	22.9	9.0
OCT 84	-	27.9	18.0

*) from Table 3.4.1.4.

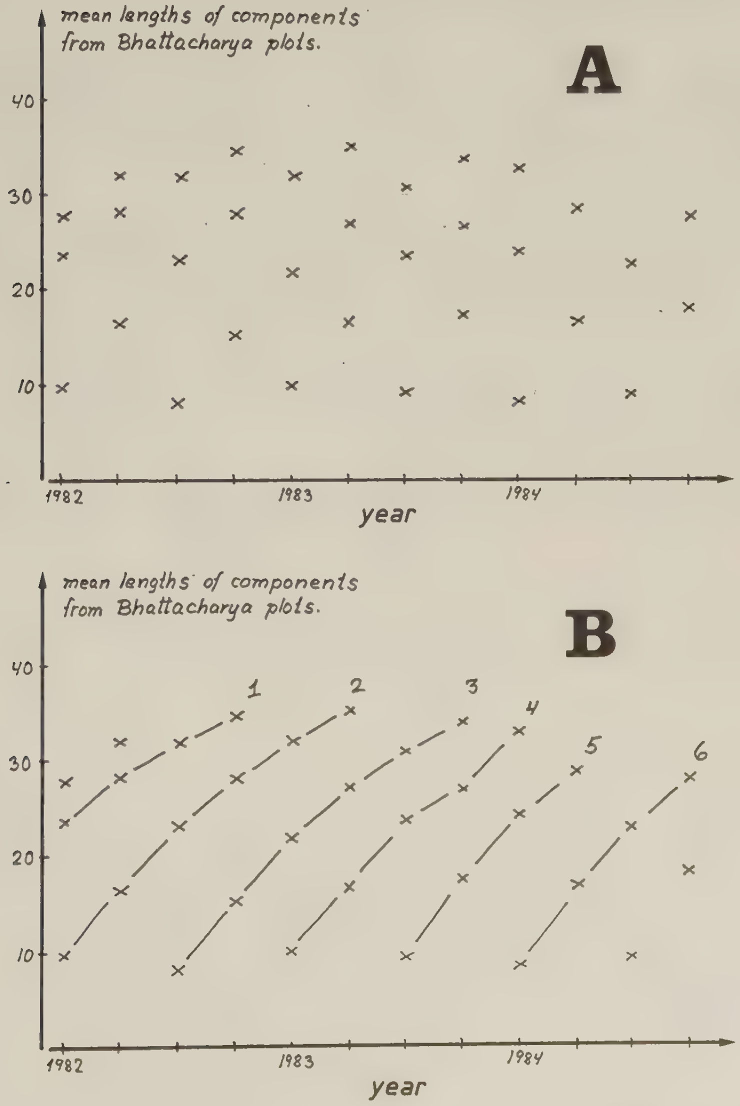


Fig. 3.4.2.1 A: Mean lengths of components from Bhattacharya plots
B: Mean lengths connected for assumed cohorts

Table 3.4.2.2 The mean lengths from Table 3.4.2.1 rearranged into cohorts (see Fig. 3.2.1.1)

date of sample	COHORTS, $\bar{L}(t)$ in cm spring 1981	autumn 1981	spring 1982	autumn 1982	spring 1983	autumn 1983
JAN 82	23.5	9.8	-	-	-	-
APR 82	28.1	16.5	-	-	-	-
JUL 82	31.8	23.1	8.0	-	-	-
OCT 82	34.6	28.0	15.3	-	-	-
JAN 83	-	32.0	21.8	10.0	-	-
APR 83	-	35.1	27.0	16.5	-	-
JUL 83	-	-	30.9	23.5	9.2	-
OCT 83	-	-	33.8	27.8	17.4	-
JAN 84	-	-	-	32.9	24.0	8.3
APR 84	-	-	-	-	28.2	16.8
JUL 84	-	-	-	-	-	22.9
OCT 84	-	-	-	-	-	27.9

Table 3.4.2.3 Input data and regression analysis for Gulland and Holt plot derived from Table 3.4.2.2

cohort	spring 1981		autumn 1981		spring 1982		autumn 1982		spring 1983		autumn 1983	
sample date	$\bar{L}$	$\frac{\Delta L}{\Delta t}$	$\bar{L}$	$\frac{\Delta L}{\Delta t}$	$\bar{L}$	$\frac{\Delta L}{\Delta t}$	$\bar{L}$	$\frac{\Delta L}{\Delta t}$	$\bar{L}$	$\frac{\Delta L}{\Delta t}$	$\bar{L}$	$\frac{\Delta L}{\Delta t}$
JAN 82												
	25.8	18.4	13.2	26.8		-		-		-		-
APR 82												
	30.0	14.8	19.8	26.4		-		-		-		-
JUL 82												
	33.2	11.2	25.6	19.6	11.7	29.2		-		-		-
OCT 82												
		-	30.0	16.0	18.6	26.0		-		-		-
JAN 83												
		-	33.6	12.4	24.4	20.8	13.3	26.0		-		-
APR 83												
		-		-	29.0	15.6	20.0	28.0		-		-
JUL 83												
		-		-	32.4	11.6	25.7	17.2	13.3	32.8		-
OCT 83												
		-		-		-	30.4	20.4	20.7	26.4		-
JAN 84												
		-		-		-		-	26.1	16.8	12.6	34.0
APR 84												
		-		-		-		-		-	20.0	24.4
JUL 84												
		-		-		-		-		-	25.4	20.0
OCT 84												

a (intercept) = 41.8361 b (slope) = - 0.8740

$$sb^2 = \frac{1}{n-2}\left((sy/sx)^2 - b^2\right) =$$

$$\frac{1}{23-2}\left((6.62333/7.08467)^2 - (-0.8740)^2\right) = 0.005242$$

sb = 0.072, t(21) = 2.09 (cf. Table 2.3.1)

95% confidence interval of b (= -K): [-1.02 , -0.72]

1983) as used in the previous section.

When each of the twelve samples is given the same treatment as the October 1983 sample the results of the twelve Bhattacharya analyses could be those given in Table 3.4.2.1. In each of the first nine samples three components have been found (as was the case in Section 3.4.1), whereas the last three samples were more difficult to analyse so that only two components could be identified.

In Fig. 3.4.2.1A the mean lengths of the components have been plotted against time whereas in Fig. 3.4.2.1B those mean lengths which we believe to correspond to the same cohorts, have been connected. Excluding the two first and the two last points we thus have identified six cohorts. The connection of points to produce cohorts is a subjective process although in the present case the choice appeared quite easy to make. In practice it may not always be so simple.

It appears from the figure that there are two cohorts per year. For instance, cohorts 3 and 4 recruited in 1982. Assuming northern seasons we call them the 1982 spring cohort and the 1982 autumn cohort. This enables us to rearrange the data of Table 3.4.2.1 into the cohorts shown in Table 3.4.2.2. The type of data in Table 3.4.2.2 makes it possible to apply the Gulland and Holt plot (cf. Section 3.3.2) by calculating:

$$\frac{\Delta L}{\Delta t} = \frac{L(t2) - L(t1)}{t2 - t1}$$

and $L(t1,t2) = 0.5 * (L(t1) + L(t2))$

where t1 and t2 stand for the dates of two consecutive lengths in the cohort time series. The time difference t2-t1 remains constant in this particular case (t2-t1 = 0.25 years). The "Exact Gulland and Holt plot" (Eq. 3.3.1.2) might therefore also have been applied.

The values of $\Delta L/\Delta t$ and L are shown in Table 3.4.2.3. To illustrate the calculations we consider the spring cohort recruited in 1981. For the two first samples we get:

$$\frac{\Delta L}{\Delta t} = \frac{L(\text{APR } 82) - L(\text{JAN } 82)}{(\text{APR } 82) - (\text{JAN } 82)} = \frac{28.1 - 23.5}{0.25} = 18.4 \quad \text{and}$$

$$L = 0.5 * (L(\text{JAN } 82) + L(\text{APR } 82)) = 0.5 * (23.5 + 28.1) = 25.8.$$

The regression of all 23 usable $\Delta L/\Delta t$ values on L values gives the results:

a (intercept) = 41.84 and b (slope) = -0.8740

from which we get K = -b = 0.87 per year and L_∞ = -a/b = 48 cm. The 95% confidence interval for K becomes [0.72 , 1.02] (cf. Table 3.4.2.3). The Gulland and Holt plot is shown in Fig. 3.4.2.2. We have thus obtained an estimate of K and L_∞ based on the entire time series.

The next step is to estimate the initial condition parameters t_{01} for the spring cohorts and t_{02} for the autumn cohorts using the von Bertalanffy plot. We allot an arbitrary age of one year to the spring cohort of 1981 in January 1982, 1.25 years in April 1982, etc. The spring cohort of 1982 is similarly allotted an age of one year in January 1983, etc. the procedure is the same for the autumn cohorts.

Table 3.4.2.4 contains the arbitrary ages, t(i) of each cohort together with the dependent variable of the von Bertalanffy plot:

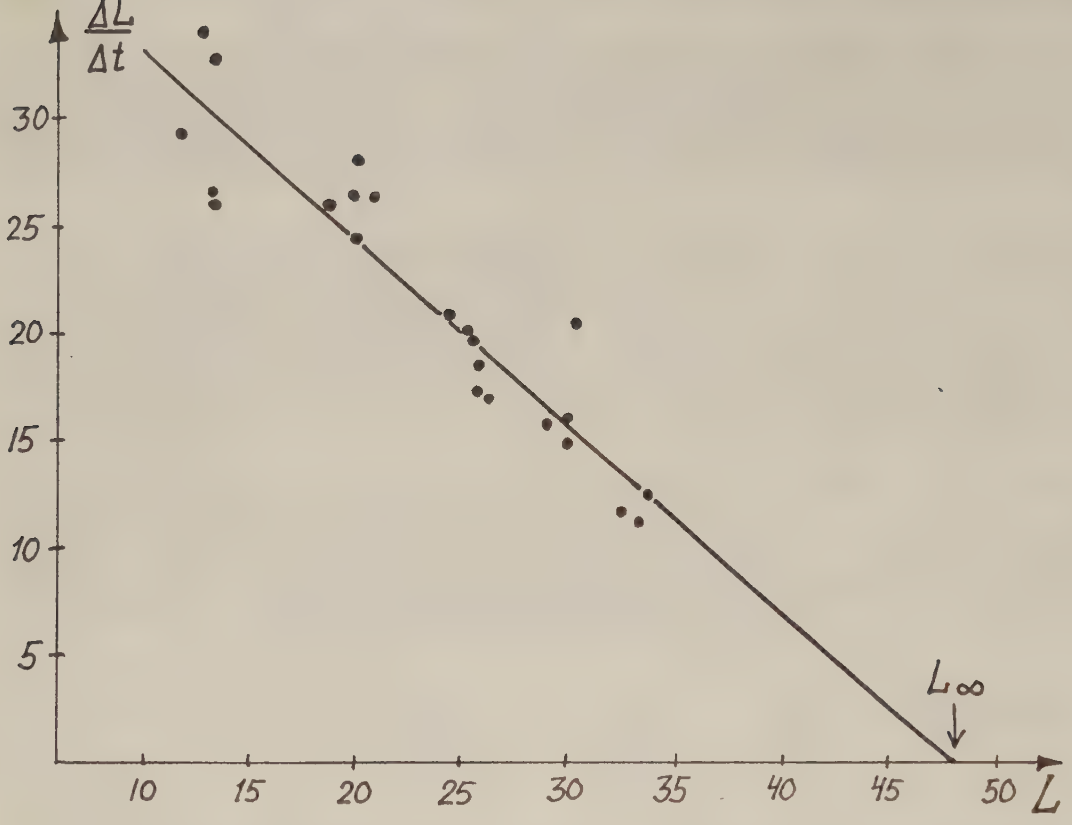


Fig. 3.4.2.2 Gulland & Holt plot based on data in Table 3.4.2.3

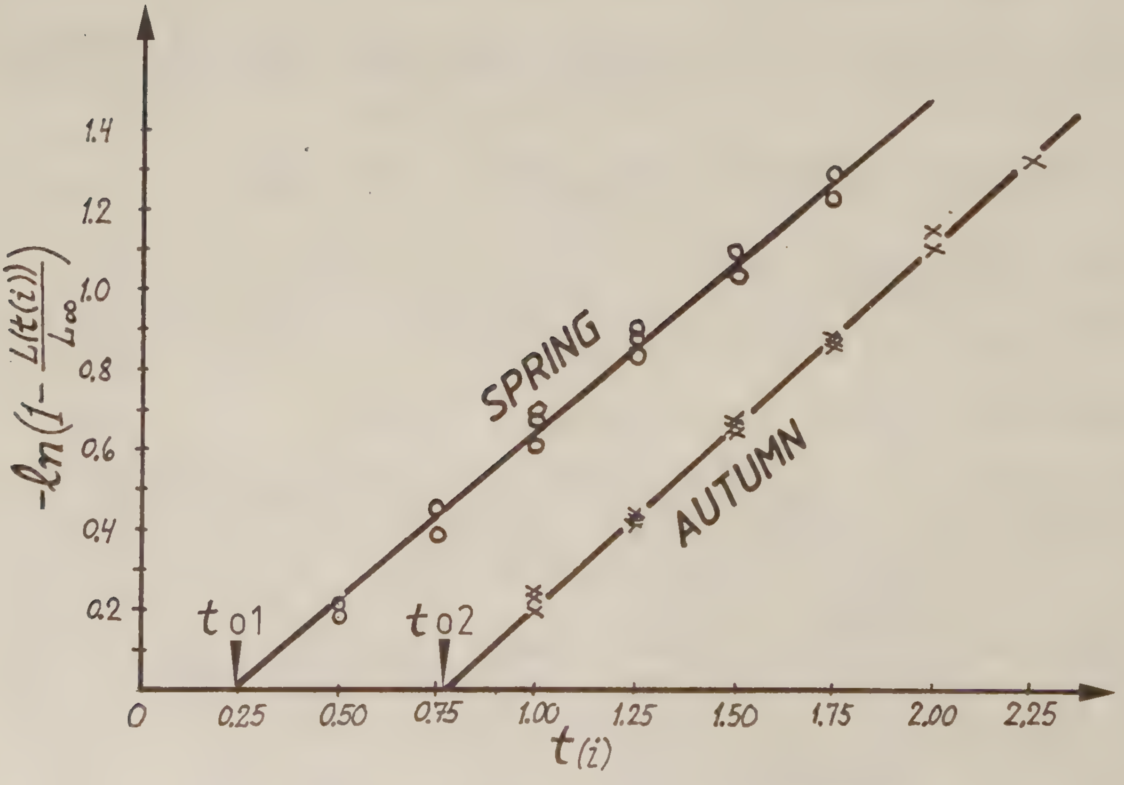


Fig. 3.4.2.3 The two von Bertalanffy plots based on data from Table 3.4.2.4

Table 3.4.2.4 Input data and regression analysis for von Bertalanffy plot $\bar{L}(t)$ derived from Table 3.4.2.2

A: spring cohorts

date of sample	spring 1981 Y(1)=1981		spring 1982 Y(3)=1982		spring 1983 Y(5)=1983		time of sampling
	t(1)	y*)	t(3)	y*)	t(5)	y*)	T = x
JAN 82	1.00	0.673	-		-		1982.00
APR 82	1.25	0.880	-		-		1982.25
JUL 82	1.50	1.086	0.50	0.182	-		1982.50
OCT 82	1.75	1.276	0.75	0.384	-		1982.75
JAN 83	-		1.00	0.605	-		1983.00
APR 83	-		1.25	0.827	-		1983.25
JUL 83	-		1.50	1.032	0.50	0.213	1983.50
OCT 83	-		1.75	1.218	0.75	0.450	1983.75
JAN 84	-		-		1.00	0.693	1984.00
APR 84	-		-		1.25	0.886	1984.25
JUL 84	-		-		-		1984.50
OCT 84	-		-		-		1984.75

spring cohorts:
a(intercept) = -0.20553, b(slope) = 0.84332, n = 14

$$sb^2 = \frac{1}{n-2}\left[(sy/sx)^2 - b^2\right] = \frac{1}{14-2}\left[(0.35016/0.41313)^2 - 0.84332^2\right]$$

$$= 0.0006005$$

sb = 0.0245, t(12) = 2.18 (cf. Table 2.3.1)
95% confidence interval of b (=K): [0.79 , 0.90]

B: autumn cohorts

date of sample	autumn 1981 Y(2)=1982		autumn 1982 Y(4)=1983		autumn 1983 Y(6)=1984		time of sampling
	t(2)	y*)	t(4)	y*)	t(6)	y*)	T = x
JAN 82	1.00	0.228	-		-		1982.00
APR 82	1.25	0.241	-		-		1982.25
JUL 82	1.50	0.656	-		-		1982.50
OCT 82	1.75	0.875	-		-		1982.75
JAN 83	2.00	1.099	1.00	0.234	-		1983.00
APR 83	2.25	1.314	1.25	0.421	-		1983.25
JUL 83	-		1.50	0.673	-		1983.50
OCT 83	-		1.75	0.866	-		1983.75
JAN 84	-		2.00	1.157	1.00	0.190	1984.00
APR 84	-		-		1.25	0.431	1984.25
JUL 84	-		-		1.50	0.648	1984.50
OCT 84	-		-		1.75	0.870	1984.75

autumn cohorts:
a(intercept) = -0.73047, b(slope) = 0.91692, n = 15

$$sb^2 = \frac{1}{n-2}\left[(sy/sx)^2 - b^2\right] = \frac{1}{15-2}\left[(0.36593/0.39491)^2 - 0.91692^2\right]$$

$$= 0.0001375$$

sb = 0.037, t(13) = 2.16 (cf. Table 2.3.1)
95% confidence interval of b (=K): [0.84 , 1.00]

*) $y = -\ln(1 - \bar{L}(t)/L_\infty)$, $L_\infty = 48.0$ cm, $\bar{L}(t)$ from Table 3.4.2.2

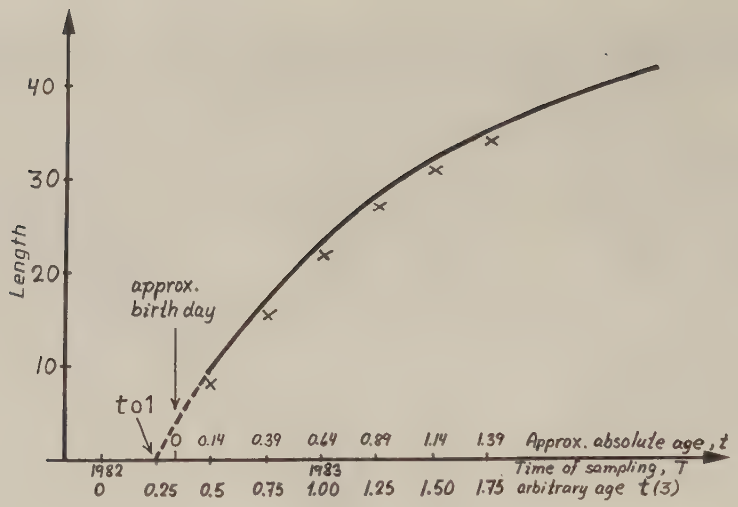


Fig. 3.4.2.4 Illustration of how the approximate birthday is estimated

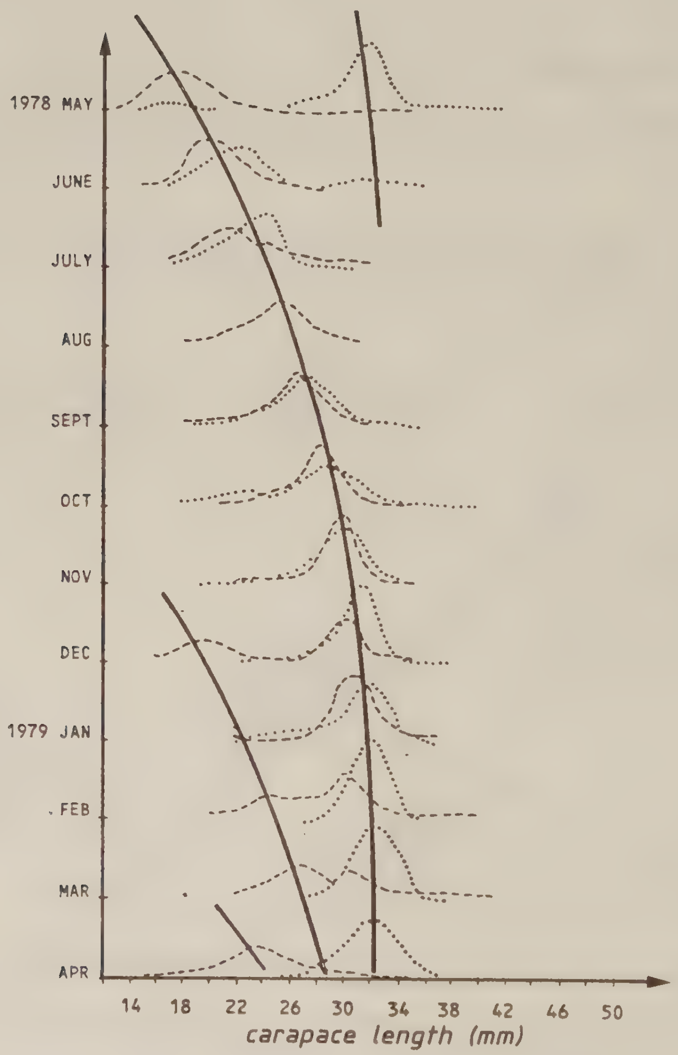


Fig. 3.4.2.5 Example of modal progression analysis. Size distributions of catches of *Penaeus semisulcatus* in the artisanal (----) and industrial (....) catches in Kuwait waters. (From Mohamed *et al*., 1979)

$$y = -\ln(1 - L(t)/L_\infty)$$

Values of L are taken from Table 3.4.2.2. There are two regression analyses to carry out:

$$\text{Spring cohorts: } y = -K\,t_{o1} + K\,t(i)\ ,\ i = 1,3,5$$
$$\text{Autumn cohorts: } y = -K\,t_{o2} + K\,t(i)\ ,\ i = 2,4,6$$

where t(i), the independent variable, is the arbitrary age of cohort no. i, as defined in Table 3.4.2.4. In this case six cohorts are considered simultaneously, and we believe that there are three spring cohorts and three autumn cohorts.

The two regression analyses gave the results:

	a (intercept)	b (slope)	$t_{o1} = - a/b$
spring cohorts:	- 0.206	0.843	0.24
autumn cohorts:	- 0.730	0.917	0.80

As expected, the difference between t_{o1} and t_{o2} became close to half a year, as explained in Section 3.2.1 (see Table 3.2.1.2) for this example. The mean of the two K-values (slopes) is 0.88 (close to the value of 0.87 estimated from the Gulland and Holt plot). A statistical test would show that these estimates are not significantly different and we would therefore use the common value K = 0.88 per year. Thus the two equations:

$$\text{Spring cohorts: } L(t) = 48 * [1 - \exp(-0.88(t - 0.24))]$$
$$\text{Autumn cohorts: } L(t) = 48 * [1 - \exp(-0.88(t - 0.80))]$$

can be used to calculate the length of spring cohorts and autumn cohorts at any time of the year.

We may stop the analysis at this stage, or we may continue trying to estimate the birthday of the cohorts. To estimate the birthday, the idea is to extrapolate the growth curve beyond the first data point and see where it intersects with the time axis as illustrated in Fig. 3.4.2.4. This figure shows cohort no. 3 as an example. The curve cuts the time axis at the point 1982.24. On the arbitrary age axis the intersection point is $t_{o1} = 0.24$. The point 1982.24 (29th of March) must be somewhere in the neighbourhood of the birthday. Because the von Bertalanffy growth curve does not conform to the early life of fish (cf. Section 3.1) this is an approximation.

An alternative way of finding the approximate birthday is to use gonodal maturity stage data, as will be discussed in Section 3.6.

The estimates obtained by following the progression of the modes (= cohorts) in the length frequencies is considered superior to the method based on one single sample (Section 3.4.1). Further, there are cases where the single sample approach is not applicable at all. This is the case for short lived species where there is only one (or are only two) cohorts in a length frequency sample. Such an example is shown in Fig. 3.4.2.5. It deals with commercial catches of the shrimp Penaeus semisulcatus in Kuwait waters (from Mohamed et al., 1979). This species has a life span of one to two years and there are two cohorts per year. Most of the samples contain only one mode, so that the single sample approach is not applicable. However, to follow the progression of the modes appears a simple thing in this case. Modal progression analysis is especially useful for such short-lived species.

The program "MODALPR" in the LFSA package (Sparre, 1987, see Chapter 14) can execute the modal progression analysis as described above. The LFSA package also allows you to continue from the Bhattacharya analysis (pro-

Table 3.4.2.5 **Illustration of the process of "data-massage". For further explanation, see text**

A: BASIC DATA: 1 cm length groups by month

Length class	1981										1982			
	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR
4- 5														
5- 6														
6- 7														
7- 8														
8- 9				18	24	12								
9-10				21	51	16								
10-11														
11-12														
12-13														
13-14														
14-15														
15-16														
16-17														
17-18														
18-19														
19-20														

B: MASSAGED DATA: 2 cm length groups by month

Length class	1981										1982			
	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR
4- 6														
6- 8														
8-10				39	75	28								
10-12														
12-14														
14-16														
16-18														
18-20														

C: MASSAGED DATA: 2 cm length groups by 3 months

Length class	1981			1982	
	MAR MAY	JUN AUG	SEP NOV	DEC FEB	MAR APR
4- 6					
6- 8					
8-10		142			
10-12					
12-14					
14-16					
16-18					
18-20					

gram "BHATTAC") with a least squares estimation of the growth parameters (program "VONBER", cf. Section 3.3.4) instead of the Gulland and Holt plot. The "COMPLEAT ELEFAN" package (Gayanilo, Soriano and Pauly, 1988) contains a program "MPA" to do the modal progression analysis. There are several other computer programs available which attempt to solve the problem dealt with in this section, some of which will be discussed in Section 3.5.

Running the Bhattacharya analysis and the modal progression analysis on a computer, one should always as a routine try out different aggregations of data, i.e. the so-called "data-massage" or "data-squeezing". Table 3.4.2.5 illustrates the process of data-massage. Part A contains the original data, i.e. a time series of 14 monthly length frequency samples grouped into 16 one cm groups. From part A to part B the data have been squeezed into 8 two cm length groups. From part B to part C data have been further squeezed into 5 three-monthly groups. Sometimes a data-massage makes the structure of the data more apparent. (With "structure" is meant the straight lines in the Bhattacharya plots and the modal progression).

If the data are grouped in such small classes that the "random noise" within each cell of the table hides the structure of the data one should massage the data. We may also observe the opposite problem, namely that the data are grouped in class intervals which are too large so that the structure becomes concealed behind the grouping. If your basic data are grouped in such large class intervals (in length or in time) there is nothing you can do to solve the problem. Therefore you should always record your basic data in as fine a grouping as practical regards allow. For example, if you are in doubt whether to use 1 cm groups or 2 cm groups, then use 1 cm groups. You can easily convert 1 cm groups into 2 cm groups, whereas you cannot do the opposite transformation. The grouping of data often simply has to be "just right" before you can carry through a successful Bhattacharya/Modal Progression Analysis.

(See **Exercise(s)** in Chapter 17).

3.4.3 The probability paper method and the parabola method

There are other ways of analysing composite normal distributions which, like the Bhattacharya analysis, are basically paper-and-pencil methods and contain a certain amount of subjectivity.

One is the probability paper method introduced by Harding (1949) and further developed by Cassie (1954). It is based on the fact that a normal distribution becomes linear when plotted on probability paper. A mixture of several normal distributions provides a more complex line with inflexion points. As with the Bhattacharya method the individual normal distributions can be removed one by one. The gist of the method is presented in Section 2.2 and Exercise 2.2.2.

Another approach is the parabola method introduced by Hald (1952) and used in fisheries research by Tanaka (1953). The mathematical base is the transformation of a normal distribution into a parabola by taking logarithms, see Section 2.6, Eq. 2.6.3. With this method, parabolas are fitted to the log-transformed numbers of composite length frequency data. The procedure is otherwise as with the Bhattacharya method which is a more sophisticated version based on the fact that differences between even-distanced points on a parabola form a straight line.

The Bhattacharya method seems to leave less to subjective decisions on the researcher's part than the other methods do. However, persons skilled in the application of either the probability paper method or the parabola method seem to reach plausible results, too.

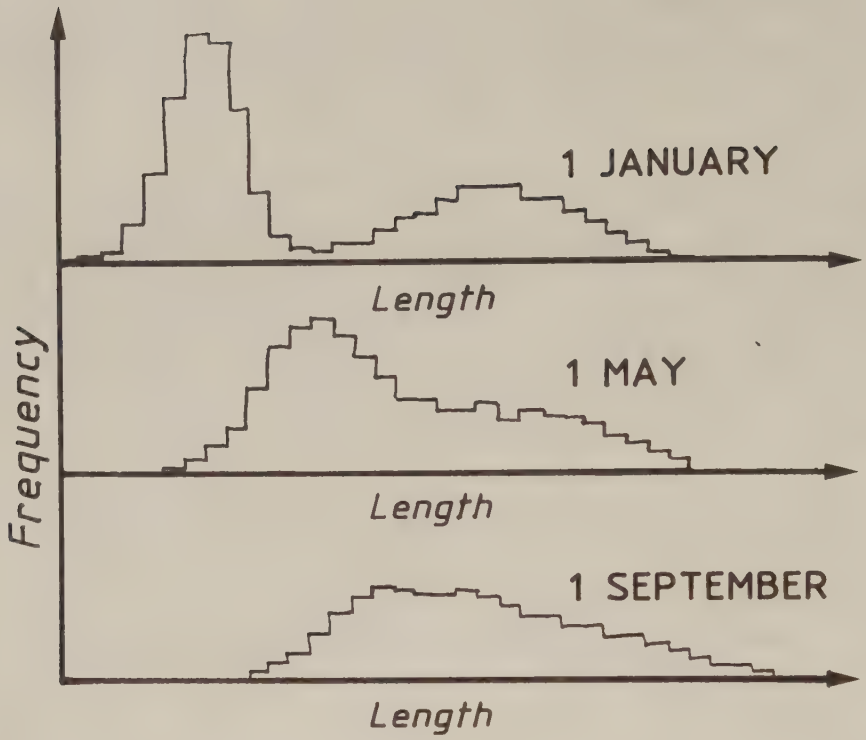


Fig. 3.5.0.1 Examples of an "easy" sample (January) and a "difficult" sample (September)

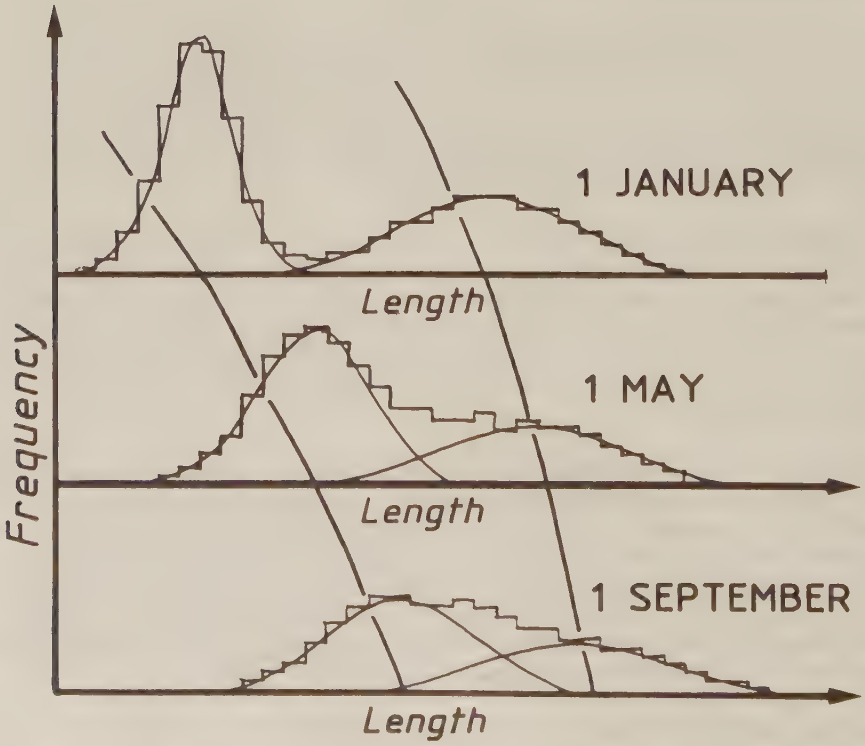


Fig. 3.5.0.2 Hypothetical example of how an "easy" sample (January) is used to treat a "difficult" sample (September)

3.5 Computer-based length frequency analysis

The methods presented in Section 3.4, the "paper and pencil" methods and their computer-based counterparts basically treat the data sample by sample. Often the tracing of the growth curves becomes easier when the entire time series is considered. Some samples may be easy to resolve into cohort components and to interpret in terms of growth in an unambiguous way. By using the findings from the "easy" samples we may also be able to give unambiguous interpretations of samples we would otherwise not be able to interpret.

Fig. 3.5.0.1 and 3.5.0.2 illustrate this feature. The January sample in Fig. 3.5.0.1 seems easy to resolve into two components as shown in Fig. 3.5.0.2, whereas the September sample shows no structure whatsoever. The May sample appears more problematic than the January sample, but still possible to interpret. Together the January and the May samples show a clear picture from which a growth curve can be estimated. Extrapolating the growth curve to the September sample we are now also in a position to split it into cohorts.

This approach may be applied when using the "paper and pencil" method, especially when aided by a computer. It is however possible to leave more work to the computer and to let the computer do the analysis using a more sophisticated technique (e.g. to apply a least squares estimation technique, cf. Section 3.3.4).

The computer-based methods to be dealt with here require so much computation that it is almost impossible to do them by paper and pencil. We present two alternative approaches:

1. The "ELEFAN I" method (Electronic LEngth Frequency ANalysis)
2. The "maximum-likelihood" method.

The first was introduced by Pauly and David (1981). The second may be considered a computerized version of the Bhattacharya method. It is based on the traditional theory on statistical analysis of frequency samples - a method which you may consider a generalized version of linear regression analysis. The basic philosophies behind the two methods are similar.

A detailed discussion of computer-based methods is considered outside the scope of the manual. The main purpose is to present some basic features of the methods which hopefully encourages the reader to go into further studies in this field.

3.5.1 ELEFAN I

The "COMPLEAT ELEFAN" is a package of microcomputer programs written in the computer language BASIC (Gayanilo, Soriano and Pauly, 1988). It contains a number of programs for fish stock assessment of which "ELEFAN I" deals with estimation of growth parameters using length frequency analysis (Pauly and David, 1981; and Pauly, 1988). The ELEFAN package is still under development and the most recent description of the entire package will be found in Pauly (1988).

EXAMPLE:

To illustrate ELEFAN I we use the data on coral trout shown in Fig. 3.4.0.2. ELEFAN I consists of two major steps:

Step 1: Restructuring of length frequencies

Step 2: Fitting of a growth curve

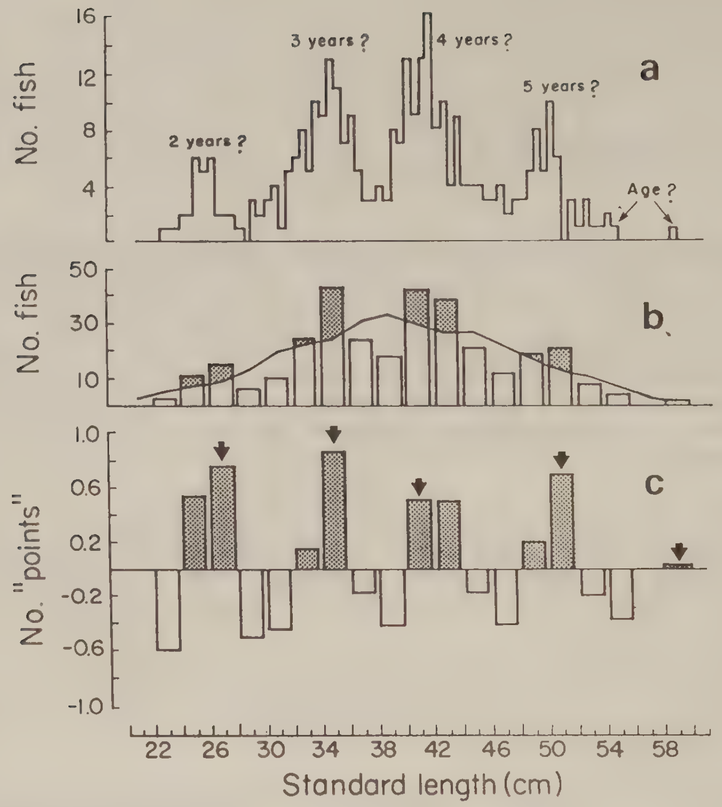


Fig. 3.5.1.1 **Example of the ELEFAN I restructuring of a length frequency sample (from Pauly & David, 1981). Data from Goeden, (1978), on the coral trout (Plectropomus leopardus)**

Step 1, the restructuring process is illustrated in Fig. 3.5.1.1 where part a shows the original data as presented by Goeden (1978) in 5 mm length groups. To smooth out small irregularities the data have been rearranged in 2 cm length groups as shown in part b. The curve in part b is the moving average frequency over 5 length groups. The method to obtain this is illustrated for the length interval 26-28 cm:

Interval	frequency
18-20	0 *
20-22	0 *
22-24	2
24-26	11
26-28	15
28-30	6
30-32	10
......	

$$\text{moving average} = \frac{2+11+15+6+10}{5} = 8.8$$

The values, for the first length groups 22-24 and 24-26 cm are calculated by adding two zeroes as indicated by "*". (A similar procedure is applied to the last length groups). This curve is used to emphasize peaks (shaded bars above moving average) and intervening throughs. Going from part b to part c the original frequencies have been divided by the moving average and 1 has been subtracted. Consider again as an example length group 26-28 cm. Here we get:

$15/8.8 - 1 = 0.7$ "points"

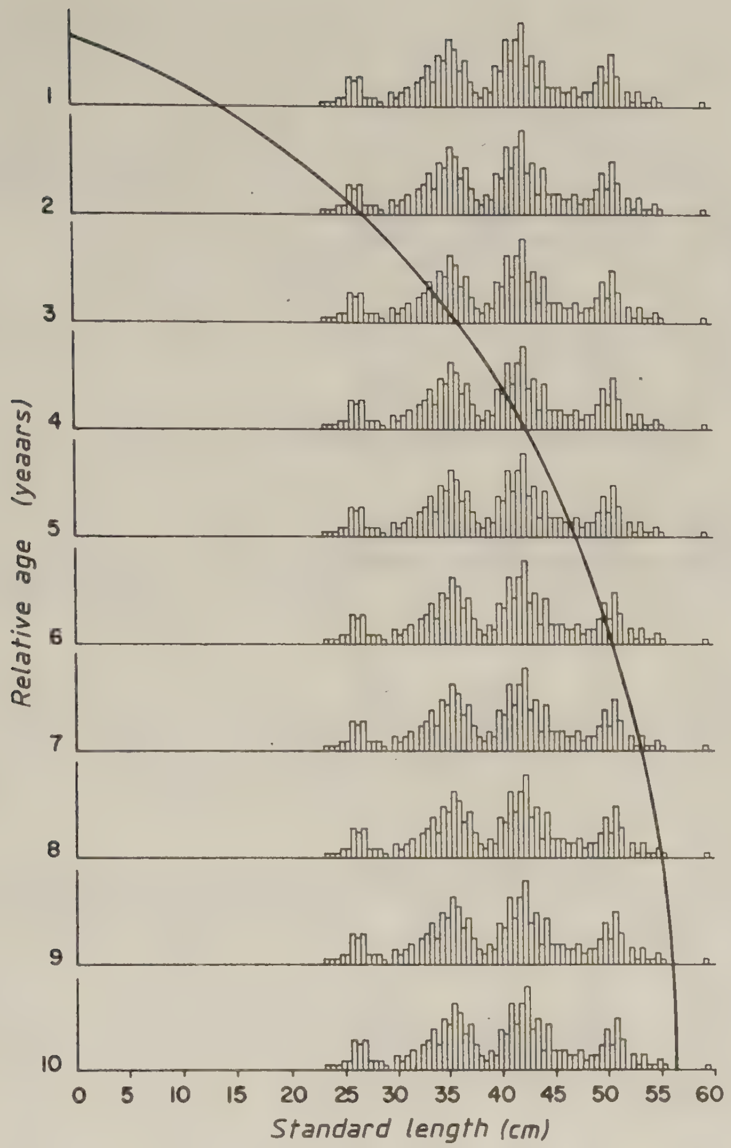
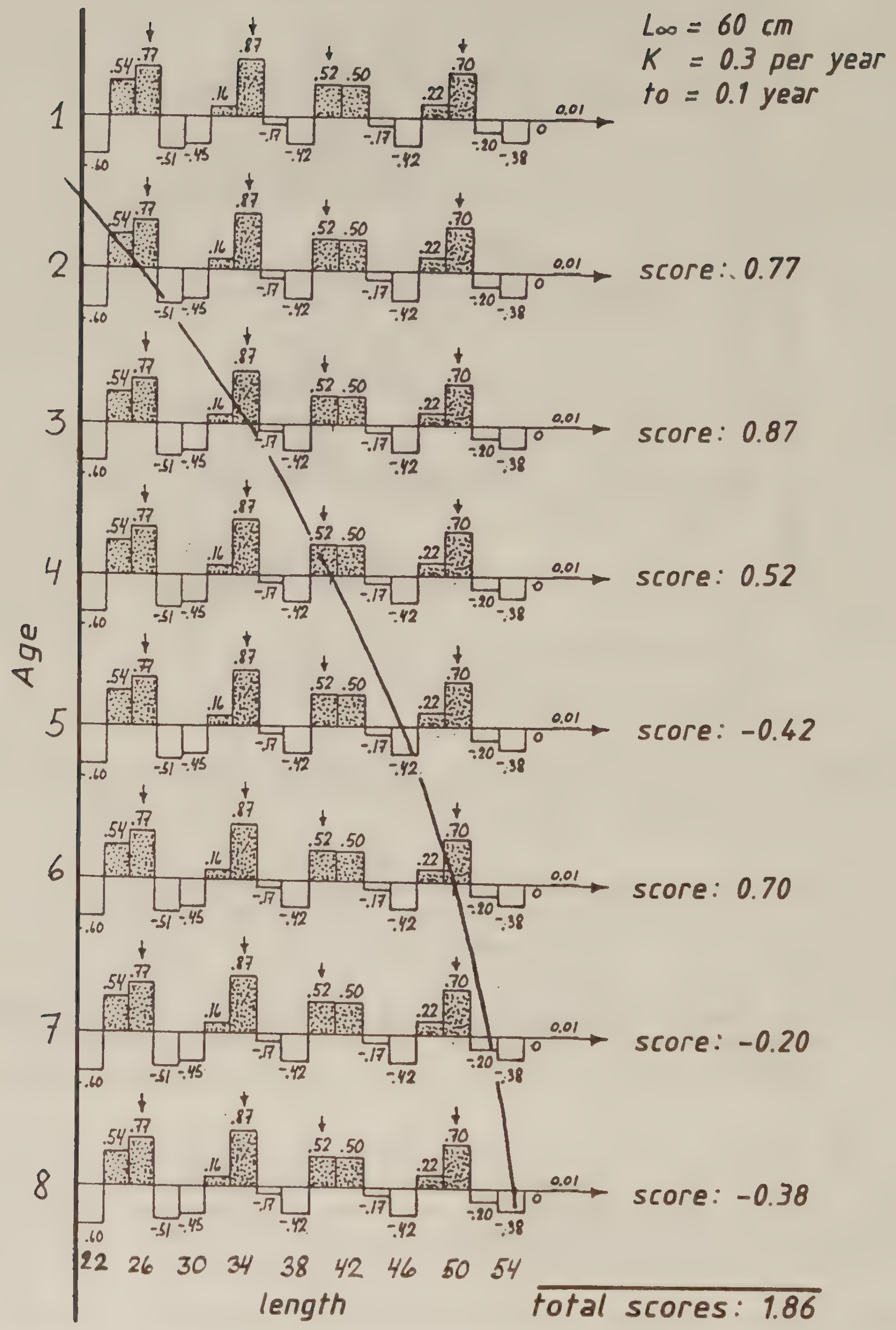


Fig. 3.5.1.2 The length frequency sample of Fig. 3.5.1.1a repeated over 10 years for simulation of time series of samples (compare with Fig. 3.5.1.3)

Actually, some additional minor adjustments have also been made but we we shall not get into that. Using the restructuring process the peaks and the throughs became well-structured and easy to identify by the "points" alotted. Note that clear peaks have been allotted a similar number of points irrespective of the number of fish they represent.

Step 2, the fitting of the growth curve is illustrated in Figs. 3.5.1.2 and 3.5.1.3.

In the present example for coral trout only one sample was used. To do the ELEFAN I type of fitting growth curves we should preferably have a time series of samples. Basically, ELEFAN I is a modal progression analysis. However, if a time series is not available we can circumvent the problem by assuming one, simply by repeating the sample for a suitable range of years, the assumption being that all cohorts follow the same growth curve. Thus, ELEFAN I applies to both the single sample case and the time series case.



$$\frac{ESP}{ASP} = \frac{1.86}{2.86} = 0.65$$

Fig. 3.5.1.3 The restructured length frequency sample of Fig. 3.5.1.1c repeated over eight years to simulate a time series of samples. (Compare Fig. 3.5.1.2). A single growth curve determined by the parameters L_∞ = 60 cm, K = 0.3 per year is tested for goodness of fit (ESP/ASP)

If the time series over the ten years shown in Fig.3.5.1.2 had been a real time series we would have got slightly different frequencies each year. Fig. 3.5.1.3 shows eight repetitions of the restructured sample arranged similarly to Fig. 3.5.1.2. It is difficult to fit a curve to the original frequencies in Fig. 3.5.1.2 and it is not possible to give an objective criterion whether one curve fits better than another if one uses an eye fit only. The restructured samples in Fig. 3.5.1.3 however, are easier to fit because peaks and throughs have been exaggerated.

With the restructured data (the "points" shown in Fig. 3.5.1.1c) it has become possible to define an objective measure for goodness of fit, for which Pauly and David (1981) suggested the ratio "ESP/ASP", where "ESP" stands for "Explained Sum of Peaks" and "ASP" for "Available Sum of Peaks".

To understand the concept of "ESP" consider Fig. 3.5.1.3. The most convincing fit of a growth curve is one which hits all the peaks indicated by arrows. However, there may not exist such a von Bertalanffy growth curve, and therefore a "score" concept has been introduced to measure how close a curve can come to the best fit. Whenever a curve hits a bar, either positive or negative it scores "points" (cf. Fig. 3.5.1.1). The total score of a growth curve is the sum of the points scored from each sample as shown in Fig. 3.5.1.3.

"ASP" (available sum of peaks) is the maximum score a curve can reach, i.e. the sum of the positive peaks indicated with arrows. Such an arrow occurs whenever there is a sequence of positive bars. (In this connection a "sequence" may be a single bar). The ratio ESP/ASP thus becomes a measure for how close a curve is to the best possible fit.

The computational procedure described so far may be carried out by paper and pencil for a single growth curve within a reasonable time. But after that it is no longer possible (in practice) to follow ELEFAN I by paper and pencil. One of the main features of ELEFAN I is that many (say, thousands) of different growth curves are tested in the way described in Fig. 3.5.1.3. Among the thousands of alternative growth curves the one that produces the highest value of ESP/ASP is selected.

The ELEFAN I approach differs from other methods discussed so far except for the von Bertalanffy plot, in the sense that it does not apply the von Bertalanffy equation in the form of Eq. 3.1.0.1:

$$L(t) = L_\infty[1-\exp(-K(t-t_o))]$$

which expresses the length as a function of age. In stead it uses the "inverse von Bertalanffy equation" (Eq. 3.3.3.2):

$$t(L) = t_o - (1/K) \ln (1 -L/L_\infty)$$

which expresses the (relative) age as a function of length. One implication of Eq. 3.3.3.2 is that L_∞ must be larger than the largest fish in the sample (cf. the discussion in Section 3.3.3). ELEFAN I therefore gives higher estimates of L_∞ than the Gulland and Holt plot for example. However, as mentioned in Section 3.4 the differences in growth curves may not be large, but one should keep this feature of ELEFAN I in mind when comparing ELEFAN I results with results obtained by other methods.

(See **Exercise(s)** in Chapter 17).

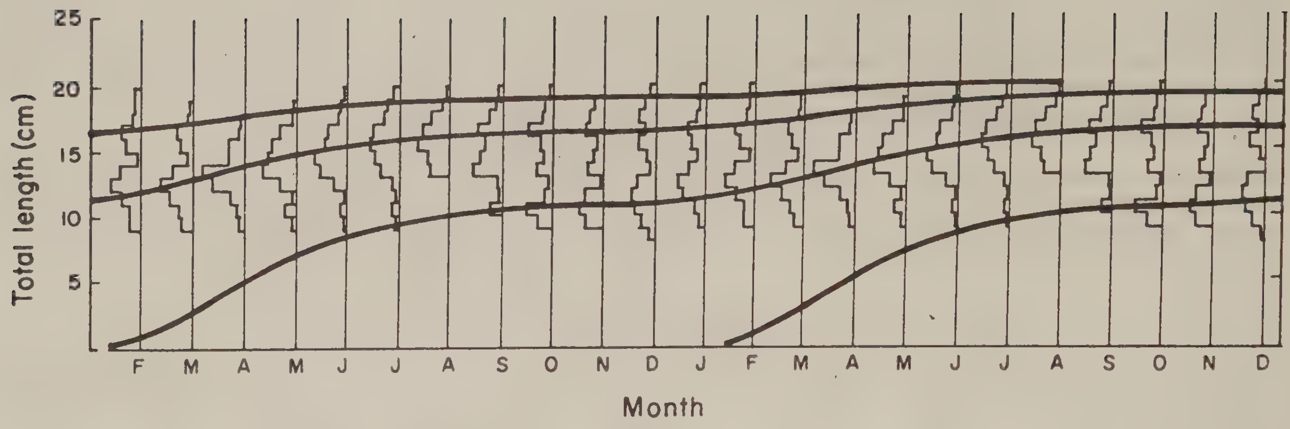


Fig. 3.5.2.1 Example of a seasonally oscillating growth curve estimated by ELEFAN I (from Pauly, 1981). Data from Rodriguez (1977) on female shrimp (Penaeus kerathurus) off Cadiz, Spain. Note that data were available for one year, and these have been repeated to simulate two years of sampling. Estimated parameters are: L_∞ = 21.0 cm (total length), K = 0.8 per year, C = 0.9, tw = 0.8 (winterpoint), ESP/ASP = 0.46

$$L(t) = L_\infty\left[1 - \exp\left\{-K(t-t_0) - \frac{CK}{2\pi}\sin 2\pi(t-t_s)\right\}\right]$$

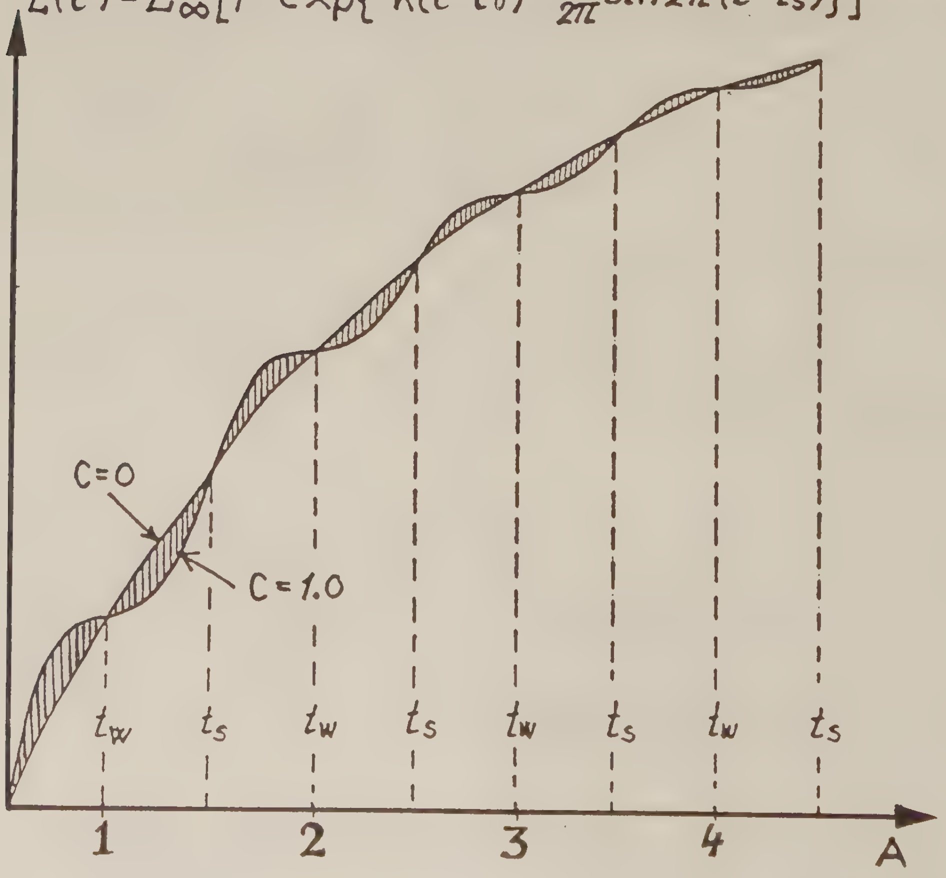


Fig. 3.5.2.2 The seasonalized von Bertalanffy growth equation. Note that for C = 1 the growth rate is zero at the winterpoints

3.5.2 The seasonalized von Bertalanffy growth equation

Fig. 3.5.2.1 shows an application of ELEFAN I to a penaeid shrimp. This growth curve estimated by ELEFAN I is clearly not a von Bertalanffy growth curve because $\Delta L/\Delta t$ does not decrease linearly with age, cf. Section 3.1. The explanation is that ELEFAN I works with the "seasonalized von Bertalanffy growth equation" (Pitcher and Macdonald, 1973; Cloern and Nichols, 1978 and Pauly and Gaschütz, 1979):

$$L(t) = L_\infty [1 - \exp(-K(t - t_o)-(CK/2\pi) \sin(2\pi(t-ts)))] \qquad (3.5.2.1)$$

This is the usual von Bertalanffy equation (Eq. 3.1.0.1) with an extra term:

$$(CK/2\pi) \sin(2\pi(t-ts)) \qquad (\text{where } \pi = 3.14159..)$$

This term produces seasonal oscillations of the growth rate, actually by changing t_o during the year. The parameter "ts" is called the "summer point" and takes the value between 0 and 1. At the time of the year when the fraction ts of the year has elapsed the growth rate is highest. At time tw = ts + 0.5, the "winter point", the growth rate is lowest. The parameter C, the "amplitude", also usually takes values between 0 and 1. If C = 0 Eq. 3.5.2.1 reduces to the ordinary von Bertalanffy equation, i.e. C = 0 implies no seasonality of the growth rate. The higher the value of C the more pronounced are the seasonal oscillations. If C = 1 the growth rate becomes zero at the winter point. Fig. 3.5.2.2 shows a seasonalized growth curve with C = 1 together with an ordinary von Bertalanffy curve (C = 0). All other seasonalized curves with different C's (but with other parameters kept constant) will be in the shaded area.

3.5.3 Maximum likelihood methods

This is a very general approach. To calculate a mean value as described in Section 2.1 and the least squares method described in Section 3.3.4 are applications of the maximum likelihood principle.

The method to be described in this section aims at solving the same problem as the ELEFAN I method and some other problems. The main difference lies in the definition of the goodness of fit. ELEFAN I uses the ratio ESP/ASP (cf. Section 3.5.1) whereas the maximum likelihood method uses the (weighted) sum of the squares of the deviations between model and observations (or measures with similar properties). In principle this measure of goodness of fit is the same as the one used in linear regression analysis (cf. Eq. 2.4.3 and Fig. 2.4.2).

The full statistical theory behind this method is complicated and so is the computer program. However, the biologist running the program does not need to know all these technical details. If the basic principles behind the method are understood few diffuculties in using the program should be encountered.

The basic idea of ELEFAN I, to follow the progression of modes and test a large number of alternative combinations of growth parameters, is also the basic idea behind the maximum likelihood approach. The measure for goodness of fit used in the maximum likelihood method is closely related to the so-called "chi-squared criterion" which is conceptually simple and therefore used in the following explanation of the method.

Suppose that we have one sample or a time series of length frequency samples, and let one of them be the one shown in Fig. 3.5.3.1. If we ordered the maximum likelihood program to fit a growth curve to that sample under the assumption that there are two cohorts represented in the sample the solution would be achieved as illustrated in Fig. 3.5.3.2.

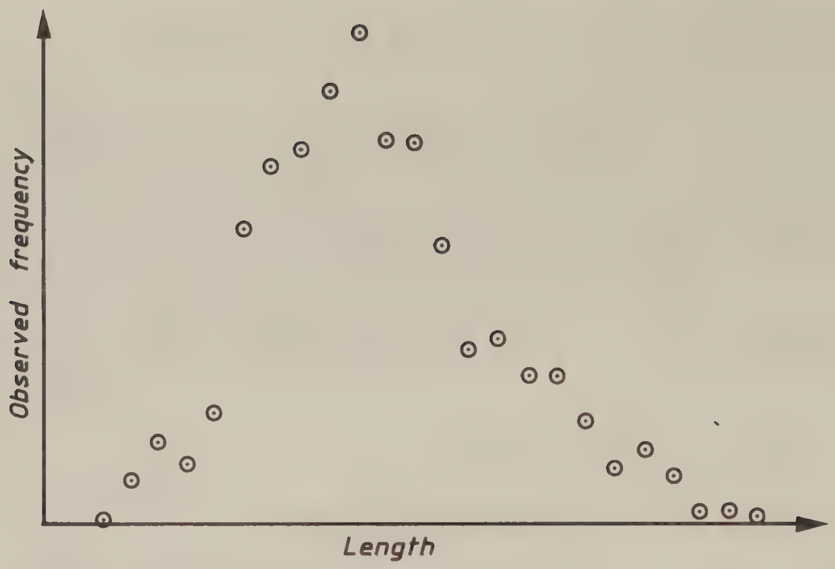


Fig. 3.5.3.1 The basic data from which the resolution into normally distributed components in Fig. 3.5.3.2 is derived

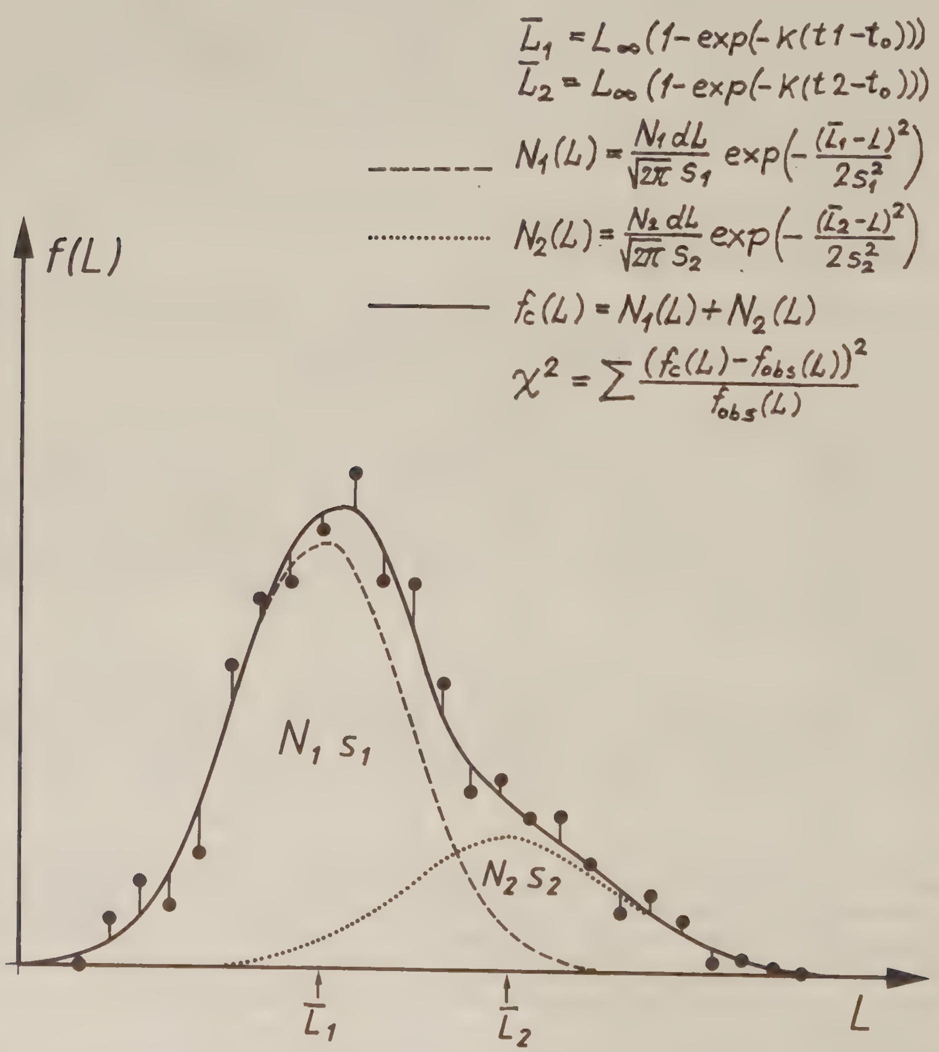


Fig. 3.5.3.2 Illustration of the chi-squared criterion. Input data are from Fig. 3.5.3.1. Also the number of cohorts must be given as input

In addition to the growth parameters the maximum likelihood method also works with the following parameters (in the case of two cohorts):

N1 = total number of observations in first cohort
N2 = total number of observations in second cohort
s1 = standard deviation of first cohort
s2 = standard deviation of second cohort

The mean lenths, $\bar{L}1$ and $\bar{L}2$ follow from the growth parameters (cf. Fig. 3.5.3.2, where $\bar{L}1$ and $\bar{L}2$ corresponding to arbitrary ages t1 and t2 are shown as an example). From the parameters the calculated (the theoretical) frequency of each cohort, N1(L) and N2(L) and the total frequency

$$fc(L) = N1(L) + N2(L)$$

of each length group can be calculated as explained in Section 2.2.

The measure of goodness of fit, the chi-squared criterion, is defined as:

$$x^2 = \Sigma \frac{(f_{obs}(L) - fc(L))^2}{f_{obs}(L)} \qquad (3.5.3.1)$$

(sum over all positive f_{obs}'s)

where $f_{obs}(L)$ stands for the observed frequency in length group L (= interval midpoint). A comparison with Eq. 2.4.3 ("fc" and "f_{obs}" correspond to "a + b x(i)" and "y(i)", respectively) illustrates the relationship between the chi-squared criterion and linear regression. Fig. 3.2.2.2 shows another example of what the maximum likelihood method would get out of a length frequency sample if it was given that the number of cohorts was six.

The maximum likelihood program determines the set of parameters (L_∞, K, t_o, N1, N2, s1 and s2) which minimize the chi-squared criterion. As this is a standard measure for goodness of fit when dealing with frequencies the list of references dealing with this concept is nearly endless. A good introduction to the theory (written for biologists) is given in Sokal and Rohlf (1981, Chapter 17).

The maximum likelihood method also gives numbers and standard deviations in addition to the growth parameters. The program requires the same input as the ELEFAN I program, but it also requires the number of cohorts in the sample as input. Often one has to guess that number. However, this extra input appears not to create great practical problems.

The maximum likelihood program works as an iterative process. That is, it must be fed an initial guess on the solution which is then improved in a number of iterative steps. Thus, to start the maximum likelihood estimation procedure we need an approximation to the solution of the exercise. Such an initial solution can be obtained from, for example, the Bhattacharya analysis and the modal progression analysis described in Sections 3.4.1 and 3.4.2. The maximum likelihood method does not make the "paper-and-pencil" methods superfluous. We still need these methods to start the iteration process and, perhaps most important, to evaluate the results. The search for an acceptable set of initial values often is the most time consuming part of the task.

Fig. 3.5.3.3 illustrates the procedure of the maximum likelihood estimation. Usually, the starting point is called the "initial guess" at the solution. However, calling it a "guess" might not be appropriate, as it has to be rather close to the final solution to make the iterative process converge. Therefore, it is important to have a simple and dependable method to get a first "good guess" at the solution. For example the Bhattacharya method and the modal progression analysis could be used.

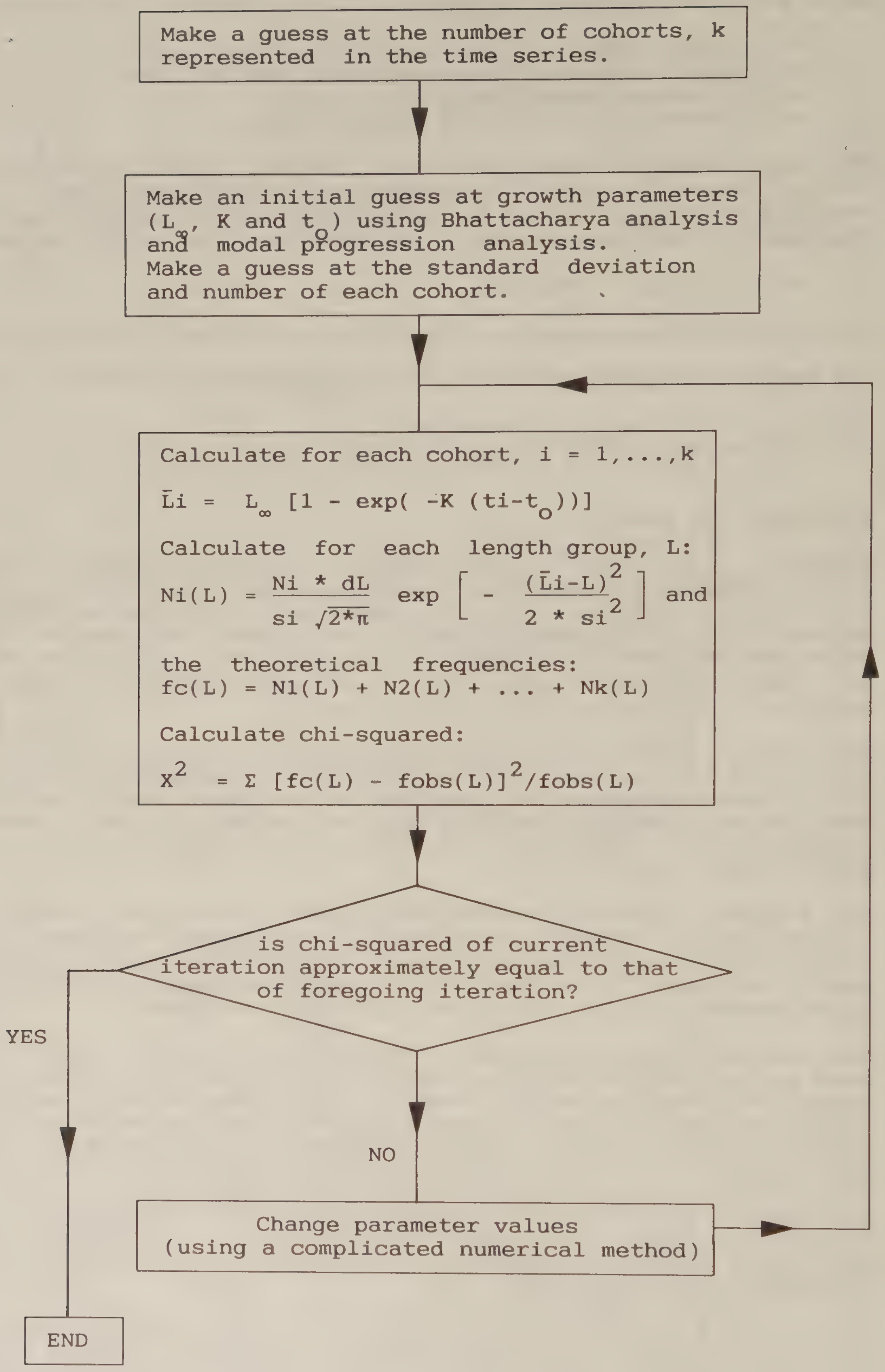

Make a guess at the number of cohorts, k represented in the time series.
Make an initial guess at growth parameters (L_∞, K and t_o) using Bhattacharya analysis and modal progression analysis.
Make a guess at the standard deviation and number of each cohort.
Calculate for each cohort, i = 1,...,k
$\bar{L}i = L_\infty [1 - \exp(-K (ti-t_o))]$
Calculate for each length group, L:
$Ni(L) = \frac{Ni * dL}{si \sqrt{2*\pi}} \exp\left[- \frac{(\bar{L}i-L)^2}{2 * si^2}\right]$ and
the theoretical frequencies:
$fc(L) = N1(L) + N2(L) + ... + Nk(L)$
Calculate chi-squared:
$X^2 = \Sigma [fc(L) - fobs(L)]^2/fobs(L)$
is chi-squared of current iteration approximately equal to that of foregoing iteration?
YES
NO
Change parameter values (using a complicated numerical method)
END

Fig. 3.5.3.3 The iterative process of the maximum likelihood estimation procedure (see also Fig. 3.5.3.2)

Another feature of the maximum likelihood method is that it gives estimates of the confidence limits of all the parameters, which the Bhattacharya method and the modal progression analysis are unable to do. The confidence limits from the modal progression analysis given in Table 3.4.2.3 are based on the assumption that the estimates from the Bhattacharya analysis have zero variance. The maximum likelihood method does not require such (highly unrealistic) assumptions.

We conclude this brief discussion of the maximum likelihood method with a few words on its historical development. The first work in the field is nearly as old as Petersen's pioneering work on length frequencies of fish (cf. Section 3.4.0) as Pearson in 1894 presented his work on separation of frequencies into normally distributed components.

One of the first computer programs to separate frequencies into normally distributed components using maximum likelihood techniques is the "NORMSEP" program by Hasselblad and Tomlinson (1971). NORMSEP was based on the work by Hasselblad (1966). Another important contribution on separation of fish length frequencies into normally distributed components was given by Macdonald and Pitcher (1979). This work was extended by Schnute and Fournier (1980) to include estimation of growth parameters in the single sample case. This contribution in turn was extended by Sparre (1987a) to deal with the time series case and the seasonalized von Bertalanffy growth curve and a few other things, the theory of which is dealt with in the following section.

3.5.4 Limitations of length frequency analysis

As appears from the examples (Section 3.4 and 3.5) it is often difficult to resolve a mixed distribution. The old fish (the longest fish) especially create problems. Intuitively, one expects the separation into components to be troublesome when mean values of neighbouring components are located close to each other compared to the size of the standard deviations.

Applying more rigorous statistical methods than those of the present manual, Hasselblad (1966), McNew and Summerfelt (1978) and Clark (1981) have shown that the "separation index"

$$I = \frac{\bar{L}(a+1) - \bar{L}(a)}{(s(a+1)+s(a))/2}$$

is a relevant quantity to study when assessing the possibility for a successful separation of two neighbour components. $\bar{L}$ stands for the mean value and s for the standard deviation (see Fig. 3.5.4.1). Without going into details the main findings of the three above mentioned works can be summarized by the rule of thumb:

> If the separation index, I, is less than two it is virtualy impossible to separate the components.

Fig. 3.5.4.1 shows an example of two normally distributed components with I = 2. Fig. 3.5.4.2 shows the typical functional relationship between separation index and variance of the estimates. (For further details see e.g. Hasselblad, 1966).

As an example consider Table 3.2.1.1 (i.e. the hypothetical data used to illustrate the paper and pencil methods). In Table 3.5.4.1 the separation indices have been calculated for the six components. These are known because the data are hypothetical or constructed. Suppose the data had been real data for which we did not know the true parameters. In that case there would be hope for estimation of only three components. This conclusion holds for all methods, including the most sophisticated computerized ones.

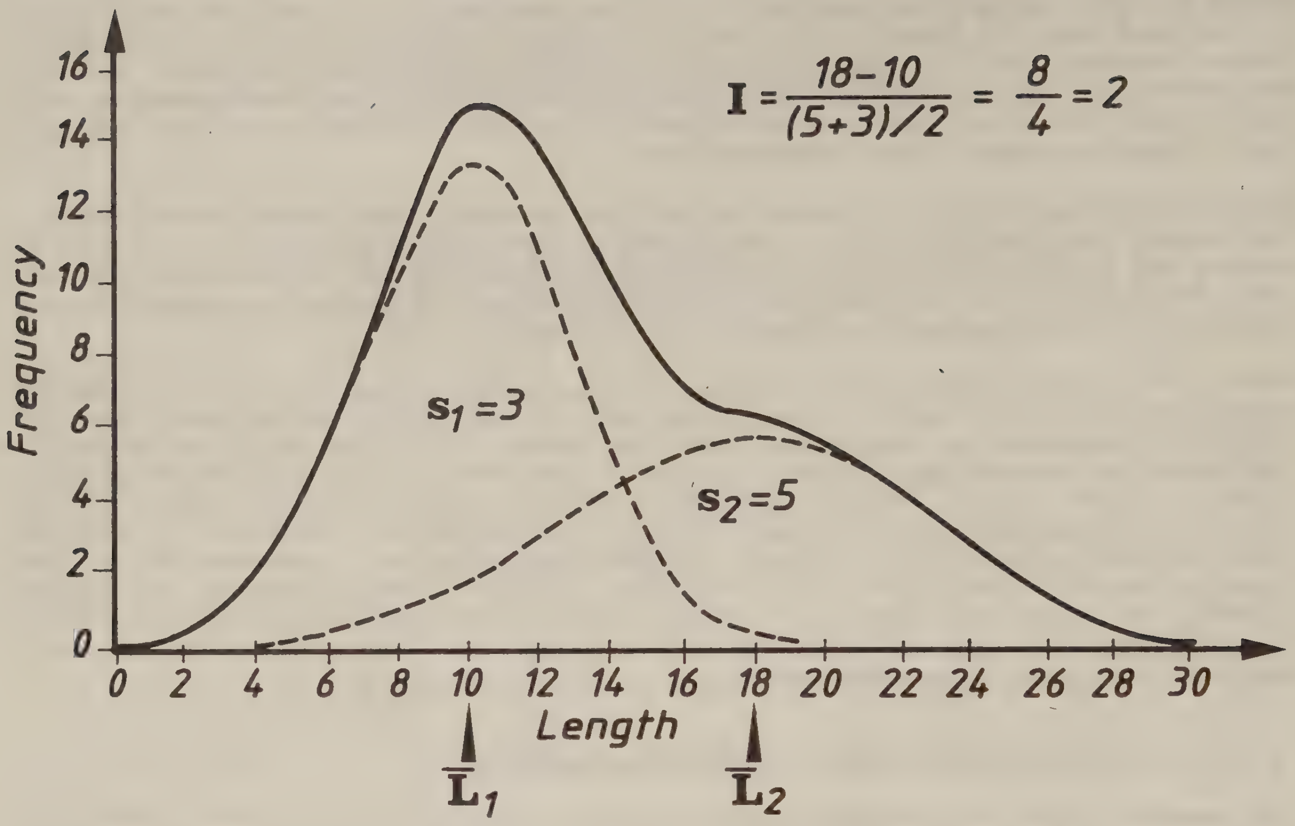


Fig. 3.5.4.1 Example of two normally distributed components with the critical separation index, I-value of 2

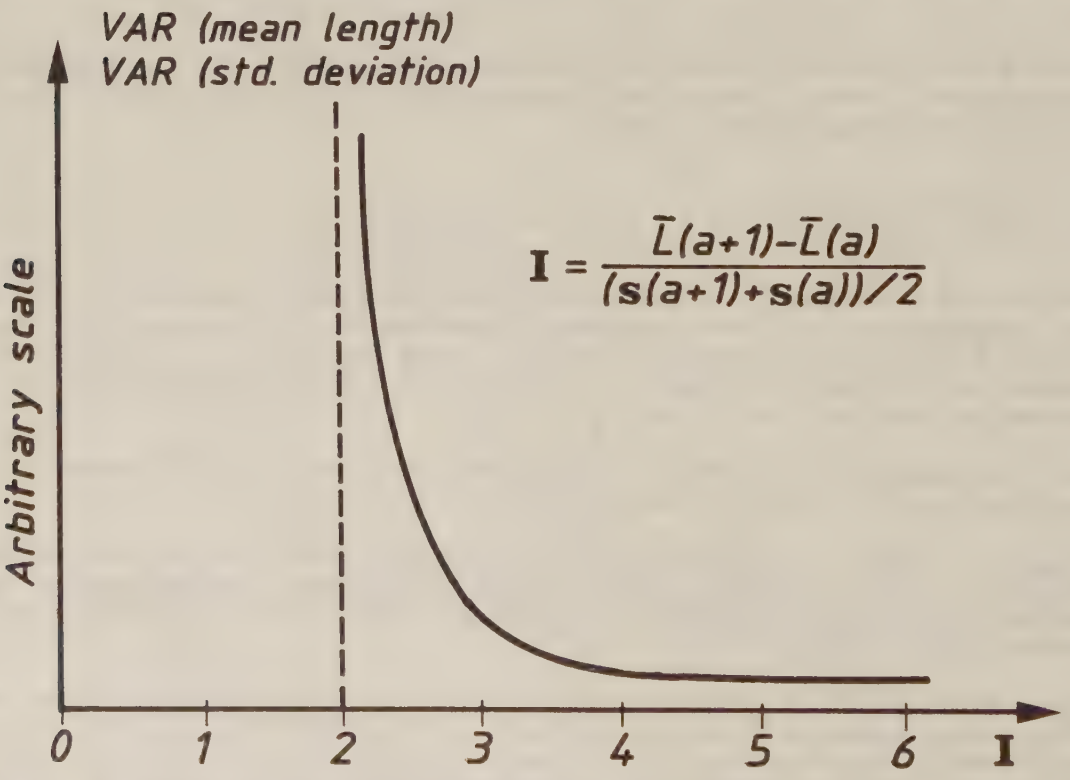


Fig. 3.5.4.2 General description of the functional relationship between separation index, I and variances of estimates

Table 3.5.4.1 Separation indices calculated for the example of Section 3.4.1. The parameters marked by "*" cannot be estimated from length frequency data alone (compare Fig. 3.2.2.2)

a	$\bar{L}(a)$	s(a)	$I = 2 \dfrac{\bar{L}(a+1) - \bar{L}(a)}{s(a+1) + s(a)}$
1	17.3	1.7	4.82
2	27.9	2.7	2.43
3	35.3	3.4	1.40
4	40.2 *	3.6 *	0.84
5	43.3 *	3.8 *	0.59
6	45.5 *	3.6 *	

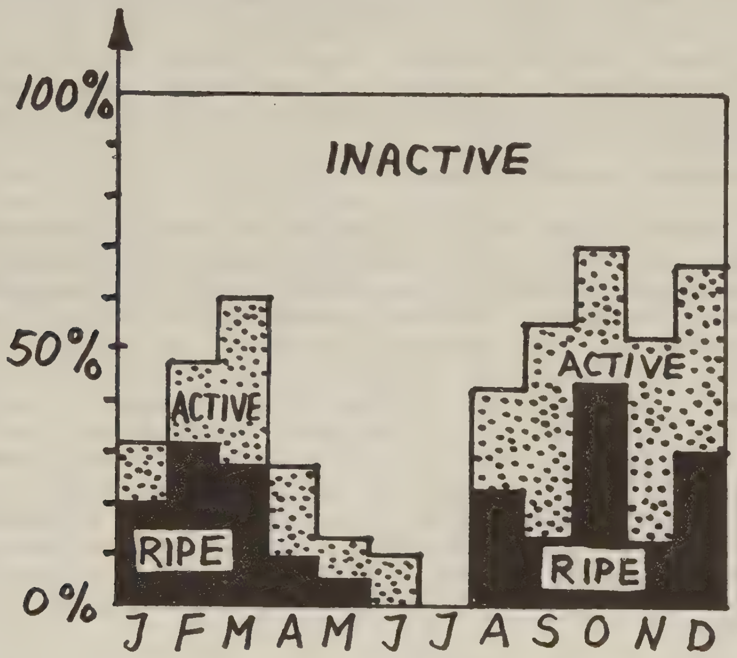


Fig. 3.6.1 Maturation stages observed for the squirrel fish (Holocentrus rufus) from Wyatt (1983). Based on samples of 1331 fish

Another way of exploring the limitations of length frequency analysis is the "Monte Carlo simulation technique". By this technique we simulate length frequency samples using a computer (cf. Section 3.2.1). The technique is called "Monte Carlo" because it includes a component of "random variability", the principle of the "roulette", which is added to all the simulated observations. By making assumptions about the parameter values and the magnitude of the random component and by simulating the corresponding length frequency samples we are in a position to evaluate the various methods. The procedure works as follows:

Step 1: Make assumptions on parameter values and the magnitude of the stochastic component.

Step 2: Simulate a time series of length frequencies according to step 1.

Step 3: Analyse the simulated data (assuming the parameters to be unknown) using for example Bhattacharya analysis and modal progression analysis.

Step 4: Compare the results (if any) of step 3 to the "true" parameters from step 1.

Using this procedure we will be able to give statements like: If a fish stock has length distributions with certain parameters then we are able to estimate/not estimate the growth parameters with a certain prespecified accuracy.

Also difficulties in obtaining unbiased samples should be mentioned in connection with the limitations of length frequency analysis. Probably the most important source of bias stems from the migration of fish. Limitations of length based methods applied to migratory fish stocks are discussed in Chapter 11.

3.6 The use of data on gonadal maturity

In Section 3.4.2 the estimation of the approximate birthday of a cohort was discussed. In that case the approximate birthday was estimated by back-tracing the modes by aid of the estimated growth curve. Another method is to estimate the spawning season from maturity stages of the adults. Fig. 3.6.1 shows an example of maturity stage data (from Wyatt, 1983). In this case three maturation graphs showing the percentages of the three main categories are presented. For a full account of stages of gonadal maturity, see Munro and Thompson (1983).

From maturation stage data, e.g., a graph of the percentage of ripe fish, we can define one (or two) mean spawning day(s), in the same way as the mean recruitment day was defined in Chapter 1, if the graph is unimodal (or bimodal). The histogram for the percentage of ripe fish in Fig. 3.6.1 could be interpreted as two spawning seasons with peaks in February and October. The mean spawning day may then be used as an estimate of the birth day (perhaps corrected for a time lag). However, the results of such analyses should be treated with a certain reservation as fluctuations in spawning is not the only factor which determines the fluctuations of recruitment. The success of a larva to feed and grow into a recruit and at the same time to avoid being eaten by predators is a complex process affected by a variety of environmental (biotic and abiotic) factors. The survival rate could, for instance, be almost nil for one spawning season and high for another. For a discussion of these matters see for example Bakun *et al*. (1982).

3.7 Comparison of growth curves, Munro's phi prime

Having estimated a set of growth parameters one would like to evaluate their reliability. The first obvious thing to do is to compare them to other growth studies on the same species (or stock) if such works exist. As discussed in Section 3.4 one should not compare the parameters one by one but compare the resultant growth curves (cf. Fig. 3.4.0.4).

Another test is the so-called "Munro's phi prime test" (Munro and Pauly, 1983, Pauly and Munro, 1984) which has gained wide application. It is suggested that the "overall growth performance" is reflected by:

$$\phi'(\text{phi prime}) = \ln K + 2 * \ln L_\infty \qquad (3.7.1)$$

(The original paper gives ϕ' in base 10 logarithms. Using natural logarithms (as done throughout this manual) gives values which are ln 10 = 2.3026 times higher than when base 10 logarithms are used). They found that species within the same family have similar ϕ'-values and with their cooperators they calculated ϕ' for a large number of species and found that within a family the ϕ's are normally distributed. For example, Moreau, Bambino and Pauly (1986) analysed 100 different tilapia populations and calculated four alternative indices for overall growth performance. They found that $\phi'$was the best index of overall growth performance, in the sense that it had minimum variance. If growth parameters are available for a number of species within the same family, it is possible to evaluate the reliability of an estimated growth curve by means of ϕ'.

The alternative indices of overall growth performance were

$$K * L_\infty, \quad \ln(K * L_\infty) \text{ and } \phi = \ln(K) + 0.67 * \ln(W_\infty)$$

where W_∞ is the asymptotic weight (Section 3.3.6). Moreau, Bambino and Pauly concluded that the first index, $K * L_\infty$, is poor (has large variance within the family) and should not be used.

Another use of Eq. 3.7.1 is to get a first rough estimate of K for a species which has not previously been studied, but which belongs to a family, some members of which have been analysed for growth parameters. If an estimate of L_∞ is available and the average ϕ' for the family members with known growth curves is available, a first estimate of K is given by:

$$K = \exp(\phi')/L_\infty^2 \tag{3.7.2}$$

which follows from solving Eq. 3.7.1 with respect to K.

Two growth curves estimated for the same stock can have the same ϕ' and yet be very different as illustrated in Fig. 3.7.1 which shows five very different growth curves, all with the same ϕ'. But if two ϕ' representing alternative estimates of growth parameters for the same stock differ greatly it indicates that one of (or both) the estimates are biased.

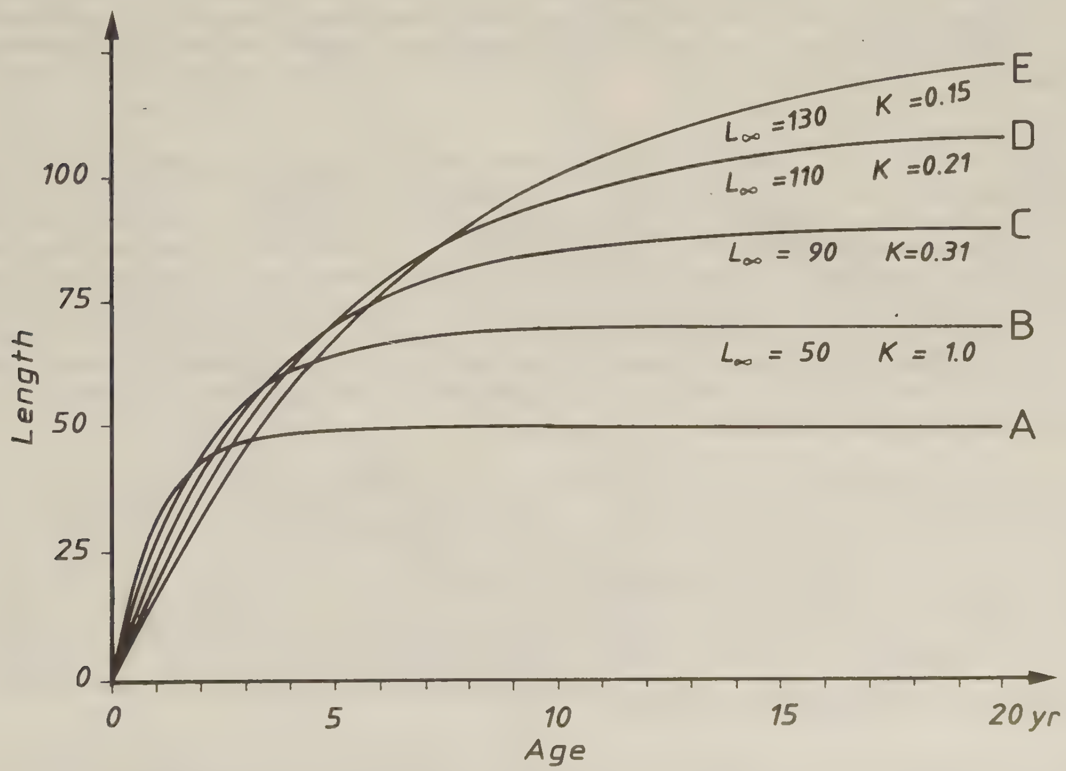


Fig. 3.7.1 Five different growth curves with the same ϕ' (phi prime)

4 ESTIMATION OF MORTALITY RATES

Chapter 3 dealt with growth, the positive aspects in the dynamics of a fish stock. This chapter deals with its negative counterpart, the death process or mortality (cf. Fig. 1.3.2). The growth was described with the aid of a model and a number of parameters, and so is the death process going to be. The key parameters used when describing death are called the "mortality rates".

This chapter deals with the definition of mortality, and it introduces some of the methods to estimate mortality rates which require input data from random samples representative of a certain part of the population, namely the exploited part. These methods do not require estimates of total catch from the population. On the other hand, they do not provide estimates of population size. Methods dealing with those aspects will be introduced in Chapter 5.

The easiest way to describe the change in numbers in a fish stock is often to follow the fate of fish spawned at approximately the same time, a cohort. We shall consider the mortality of a cohort as composed of the mortality caused by fishing and that due to all other causes lumped together as the "natural mortality". The latter covers events such as predation, disease and deaths due to old age.

4.1 The concept of a cohort and some basic notation

A "cohort" is a batch of fish all of approximately the same age and belonging to the same stock (cf. Section 1.3.1). The concept of "day of recruitment of a cohort" and "recruitment" were introduced in Section 1.6. In all the following derivations we assume (with Beverton and Holt, 1957) that a cohort consists of "average fish" only. This means that all fish of a cohort are assumed to have the same age at a given time so that they all attain the "recruitment age", Tr (cf. Section 1.6) at the same time. In Chapter 3 we similarly used the average length of a cohort to describe growth. In the context of mortality rates we are interested in the number of survivors from a cohort as a function of time (cf. Fig. 1.3.2A).

The symbol "N(t)" is used to designate the "number of survivors from a cohort" attaining age t. The age is usually measured in units of years. Thus, N(Tr) is the "number of recruits" to the fishery. Often the symbol "R" is used to designate the "recruitment", R = N(Tr). The actual choice of Tr is not critical, since all calculations are based on relative ages (age differences). In many applications we do not need to define the recruitment age. Tr is the minimum age at which the fish can enter the fishery, i.e. become liable to encounter with fishing gears (Beverton and Holt, 1957). The age at which they actually enter the fishery, Tc, is dependent on the mesh size.

Fig. 4.1.1 illustrates the basic features of cohort dynamics. Due to mortality (fishing or natural causes) there is a continuous decrease in the number of survivors. At birth the cohort has age zero. From age 0 to Tr the cohort is in the "pre-recruitment phase". In the present context we are not concerned about what happens before age Tr. After age Tr the fish may be caught if a suitable small meshed gear is used. At age Tc the fish start to be caught with the mesh size actually in use. By definition we must have Tr ≤ Tc. Tc is called the "age at first capture" and marks the beginning of the "exploited phase".

In some applications we consider several cohorts at a time as illustrated by Fig. 4.1.2, where a situation with two cohorts per year during a period of two years is depicted. Fig. 4.1.2 is a parallel to the multi cohort

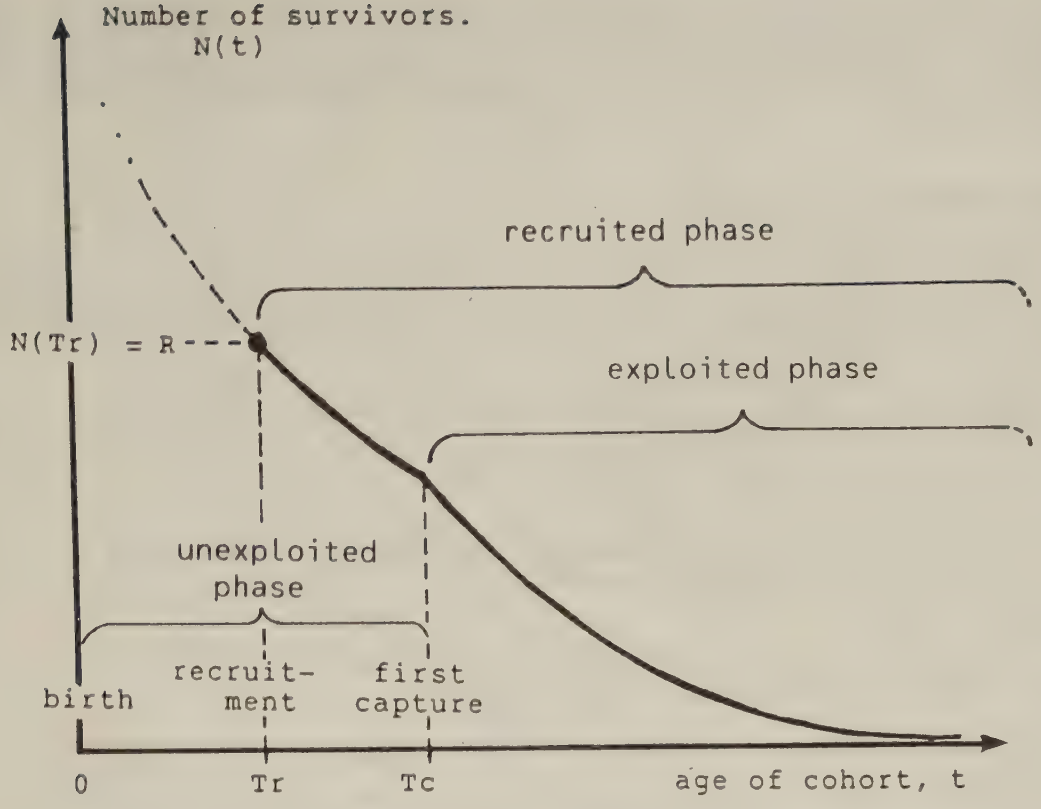

Number of survivors.
N(t)
recruited phase
N(Tr) = R
exploited phase
unexploited phase
birth
recruit-ment
first capture
0
Tr
Tc
age of cohort, t

Fig. 4.1.1 Basic features of cohort dynamics

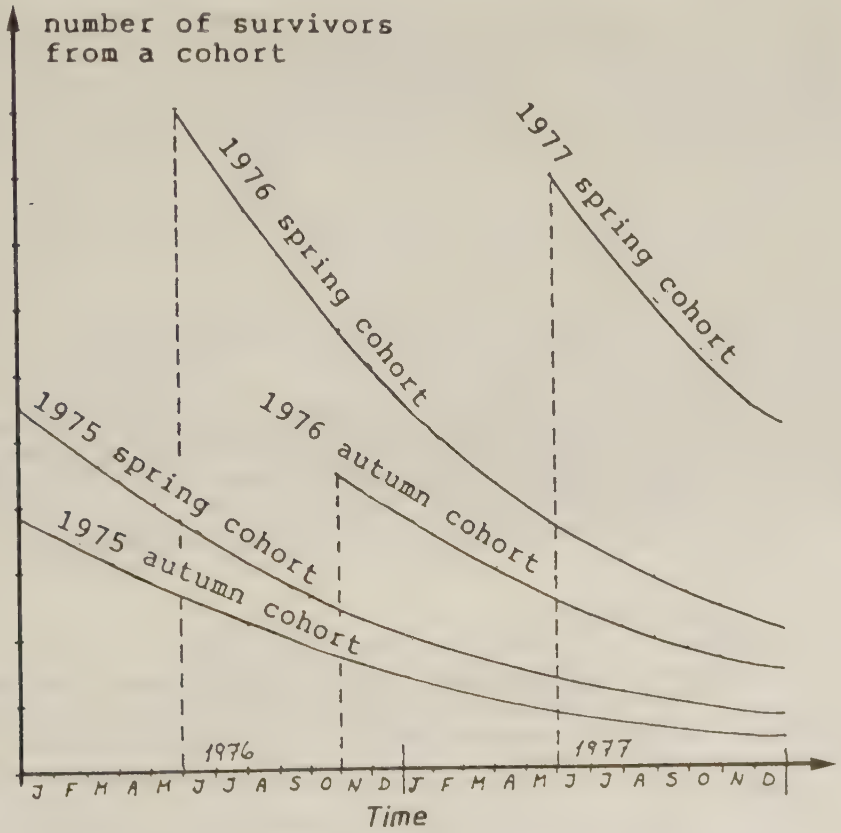

number of survivors from a cohort
1976 spring cohort
1977 spring cohort
1976 autumn cohort
1975 spring cohort
1975 autumn cohort
1976
1977
J F M A M J J A S O N D J F M A M J J A S O N D
Time

Fig. 4.1.2 Illustration of the situation with more than one cohort (compare Fig. 3.2.1.1)

case for growth which was illustrated in Fig. 3.2.1.1. If more than one cohort is considered at a time, the symbol $N(t)$ is not sufficient, and some index to indicate the cohort is then required. However, in the following only one cohort at a time is usually considered, so the short symbol $N(t)$ will do.

4.2 The dynamics of a cohort

As an example, consider the number of survivors at age $t = 0.5$ years, $N(0.5)$ and the number of survivors one day later, $N(0.50274)$ (1 day = 1/365 years = 0.00274 years). The number of specimens lost during that day is: $N(0.5) - N(0.50274)$

To designate the change in numbers during a relatively short time period (say, 1 day) we use the symbol ΔN:

$$\Delta N(0.5) = N(0.50274) - N(0.5)$$

Note that ΔN is negative because it represents a loss from the cohort. The rate of change in numbers is written:

$$\Delta N(t) / \Delta t$$

where Δt is the length of the time period.

Suppose $N(0.5) = 1000$ and $N(0.50274) = 997$, then

$$\frac{\Delta N(0.5)}{\Delta t} = \frac{997 - 1000}{0.00274} = -\ 1095 \text{ per year}$$

Obviously, the rate of change in numbers per year, $\Delta N/\Delta t$, depends on the number of survivors, N. The more survivors there are, the more will die. The high number of deaths (more than $N(0.5)$!) is due to the fact that it is calculated as if we every day started with $N(0.5) = 1000$ fish. $\Delta N/\Delta t$ is not dependent on Δt, as a reduction of Δt will reduce ΔN accordingly. Thus it is natural to assume $\Delta N/\Delta t$ to be proportional to N:

$$\frac{\Delta N(t)}{\Delta t} = -\ Z\ N(t) \qquad (4.2.1)$$

For the example above: $\dfrac{\Delta N(0.5)}{\Delta t} = -\ Z\ N(0.5)$ or $\dfrac{-3}{0.00274} = -1.095 * 1000$,

the coefficient of proportionality, Z, becomes 1.095 per year. Z is called the "instantaneous rate of total mortality", the "total mortality coefficient" or simply the "total mortality rate". The unit of Z is "per year", or in general "per time unit". If Z remains constant throughout the life of the cohort it can be shown that Eq. 4.2.1 is mathematically equivalent to:

$$N(t) = N(Tr)\exp(\ -Z(t-Tr)) \qquad (4.2.2)$$

Eq. 4.2.2 is called the "exponential decay model". and (together with the growth equation) it is a corner-stone of the theory of exploited fish stocks (cf. Baranov, 1918; Thompson and Bell, 1934; Fry, 1949 and Beverton and Holt, 1957). Fig. 4.2.1 shows a family of exponential decay curves for different Z-values).

To illustrate the concept of mortality coefficient we may calculate the corresponding percentage of survivors, say, 1 year after recruitment and 2 years after recruitment for $Z = 0.5$ and $Z = 2.0$ (compare Fig. 4.2.1).

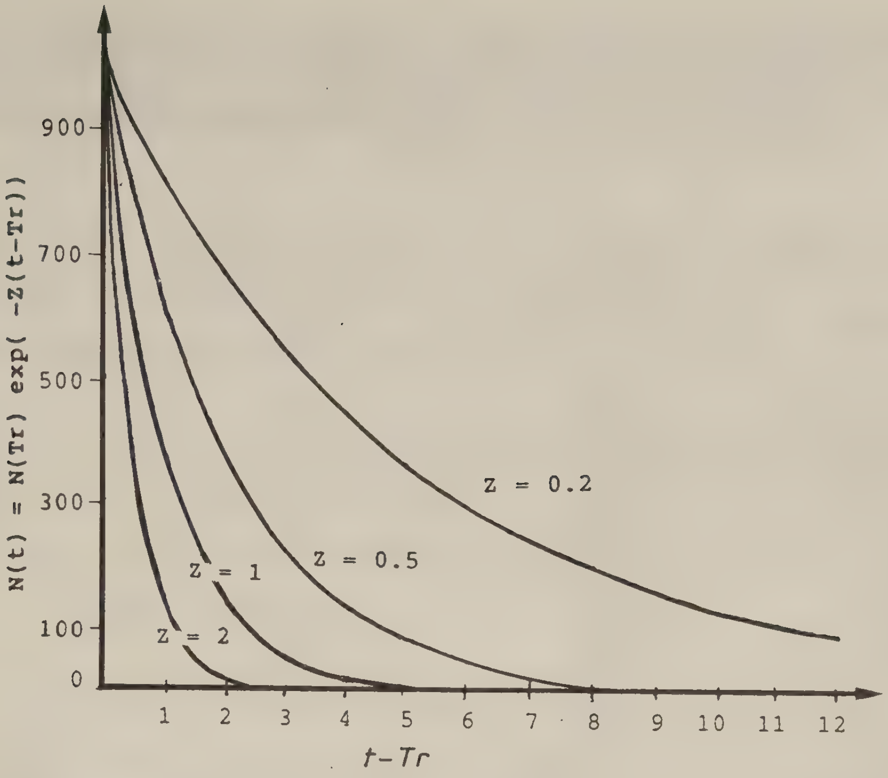


Fig. 4.2.1 **Exponential decay curves, for Z = 0.2, 0.5, 1 and 2 per year, with recruitment, N(Tr) = 1000 fish**

total mortality rate Z	percentage of survivors *) after 1 year 100 N(1+Tr)/N(Tr)	after 2 years 100 N(2+Tr)/N(Tr)
0.5	61%	37%
2.0	14%	2%

*) Percentage of survivors = 100 exp(-Z(t-Tr)), and t-Tr = 1 and 2 respectively

Table 4.2.1 shows an example of exponential decay, and it is demonstrated that the model fulfills Eq. 4.2.1 for different values of t. Fig. 4.2.2 shows the corresponding decay curve. In order to simplify the example we have here assumed that Tr = Tc, i.e. that the fish start to be caught as soon as they become available to the fishery (cf. Fig. 4.1.1). The fraction surviving after one year is called the "survival rate" (Ricker, 1975):

$$\exp(-Z) = \frac{N(tr+1)}{N(tr)}$$

Let t1 and t2 be two ages with t1 < t2. The total number of deaths during the time interval from t1 to t2 is

$$N(t1) - N(t2)$$

An example is given in Table 4.2.1 and Fig. 4.2.2 where t1 = Tr + 0.4 and t2 = Tr + 0.8 years. The number of deaths over this period is the difference between the number of survivors at t1 and t2:

$$54881 - 30119 = 24762$$

Table 4.2.1 Exponential decay for a cohort with recruitment N(Tr) = 100000 and total (constant) mortality Z = 1.5 per year. It is demonstrated that the equation /\N//\t = - Z N is fulfilled for various values of t. The corresponding graph is shown in Fig. 4.2.2. /\t is one day

age of cohort t years	number of survivors N(Tr)exp(-Z(t-Tr))	change in numbers during one day, /\N	$-\frac{/\backslash N}{/\backslash t}\,\frac{1}{N} = Z$
Tr	100000		
Tr+0.0+/\t	99590	-410	1.5
Tr+0.2	74081		
Tr+0.4	54881		
Tr+0.6	40657		
Tr+0.6+/\t	40490	-167	1.5
Tr+0.8	30119		
Tr+1.0	22313		
Tr+1.5	10540		
Tr+1.5+/\t	10497	-43	1.5
Tr+2.0	4978		
Tr+2.5	2351		
Tr+3.0	1111		
Tr+3.5	524.75		
Tr+3.5+/\t	522.60	-2.15	1.5
Tr+4.0	247.88		
Tr+5.0	55.31		
Tr+8.0	0.61		

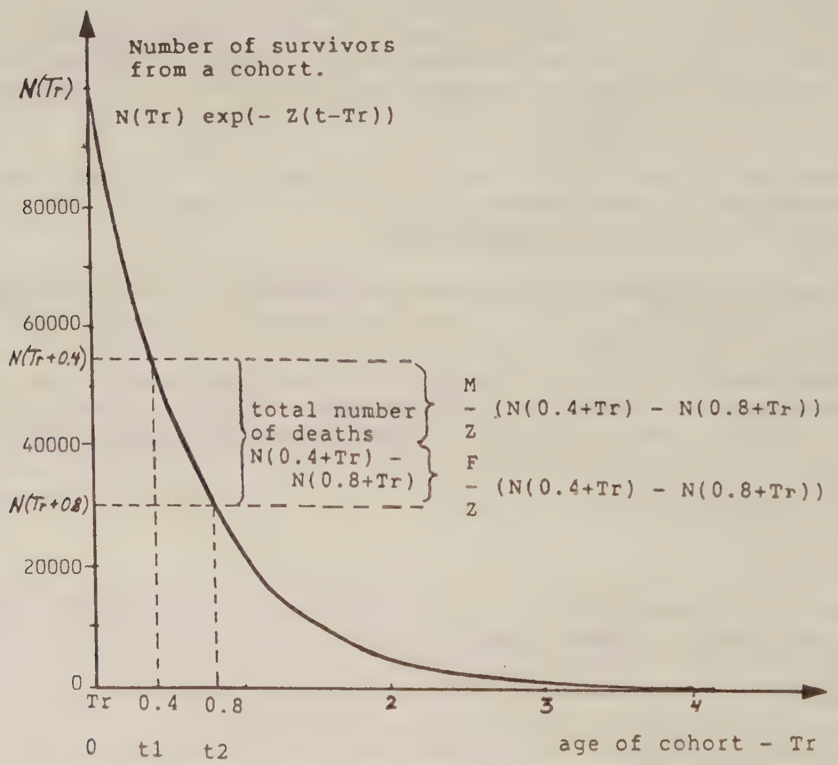


Fig. 4.2.2 Exponential decay curve with N(Tr) = 100,000 recruits and Z = 1.5 (data from Table 4.2.1). Z remains constant for all ages, t.

Of these 24762 fish some died due to fishing and the rest died due to other causes (natural deaths). To quantify this difference the "fishing mortality coefficient", F, and the "natural mortality coefficient", M, are introduced. As Z is the total mortality rate it follows that

$$Z = F + M \tag{4.2.3}$$

The deaths can thus be split into a number dying due to fishing (C) and a number dying due to natural causes (D). The number dying due to fishing during the time period from t1 to t2 is:

$$C(t1,t2) = \frac{F}{Z}\,(N(t1)-N(t2)) \tag{4.2.4}$$

Eq. 4.2.4 is called the "catch equation" (also called "Baranov's equation"), under the assumption of constant F and M (and thus constant Z) between ages t1 and t2. (Baranov, 1918).

The number dying due to natural causes correspondingly is:

$$D(t1,t2) = \frac{M}{Z}\,(N(t1)-N(t2)) \tag{4.2.5}$$

Note that $N(t1)-N(t2) = C(t1,t2) + D(t1,t2)$ because $F/Z + M/Z = 1.0$.

The fraction of deaths caused by fishing, F/Z, is called the "exploitation rate", E. Fig. 4.2.2 gives an example with Z = 1.5, F = 0.6 and M = 0.9 (compare with Table 4.2.1, where Z = 1.5). For t1 = Tr + 0.4 and t2 = Tr + 0.8 the numbers are:

$$C(Tr+0.4,Tr+0.8) = (0.6/1.5) * (54881 - 30119) = 9905$$

$$D(Tr+0.4,Tr+0.8) = (0.9/1.5) * (54881 - 30119) = 14857$$

$$N(Tr+0.4)-N(Tr+0.8) = 54881 - 30119 = 24762$$

In real life the mortality usually varies with the age of the cohort. Small (young) fish are exposed to a greater natural mortality because more predators can eat them. But small fish may on the other hand suffer less fishing mortality than large (old) fish because they either have not yet migrated to the fishing grounds or they escape through the meshes of the gear. However, if the time span from t1 to t2 is not too large it is considered a fair approximation to assume F and M to remain constant within this period.

The "catch equation" (4.2.4) is one of the most important mathematical expressions in fisheries biology. For many applications, however, it is convenient to rearrange its terms. To do so Eq. 4.2.2 is applied to N(t2):

$N(t2) = N(Tr)\exp(-Z(t2-Tr))$ which is equivalent to

$N(t2) = N(Tr)\exp(-Z(t1-Tr))\exp[-Z(t2-t1)]$ and to

$$N(t2) = N(t1)\exp[-Z(t2-t1)] \tag{4.2.6}$$

Inserting Eq. 4.2.6 into Eq. 4.2.4 gives:

$$C(t1,t2) = N(t1)\,\frac{F}{Z}\,[1 - \exp(-Z(t2-t1))] \tag{4.2.7}$$

Eq. 4.2.7 is the most commonly used version of the catch equation. Another is convenient for special applications:

$$C(t1,t2) = (t2-t1)F\ \bar{N}(t1,t2) \qquad (4.2.8)$$

where $\bar{N}(t1,t2)$ is the "average number of survivors during the time period from t1 to t2". To get consistency between Eq. 4.2.7 and Eq. 4.2.8 we must have:

$$\bar{N}(t1,t2) = N(t1)\ \frac{1}{Z\ (t2-t1)}\ [1 - \exp(-Z(t2-t1))] \qquad (4.2.9)$$

The mathematical proof of Eq. 4.2.9 is outside the scope of this manual. However, in Fig. 4.2.3 it is demonstrated that Eq. 4.2.9 conforms to the intuitive concept of "average number". For the example used above (Fig. 4.2.2 and Table 4.2.1) we find:

$$\bar{N}(Tr+0.4,Tr+0.8) = 54881\ \frac{1}{1.5 * (0.8-0.4)}\ [1-\exp(-1.5 * (0.8-0.4))]$$

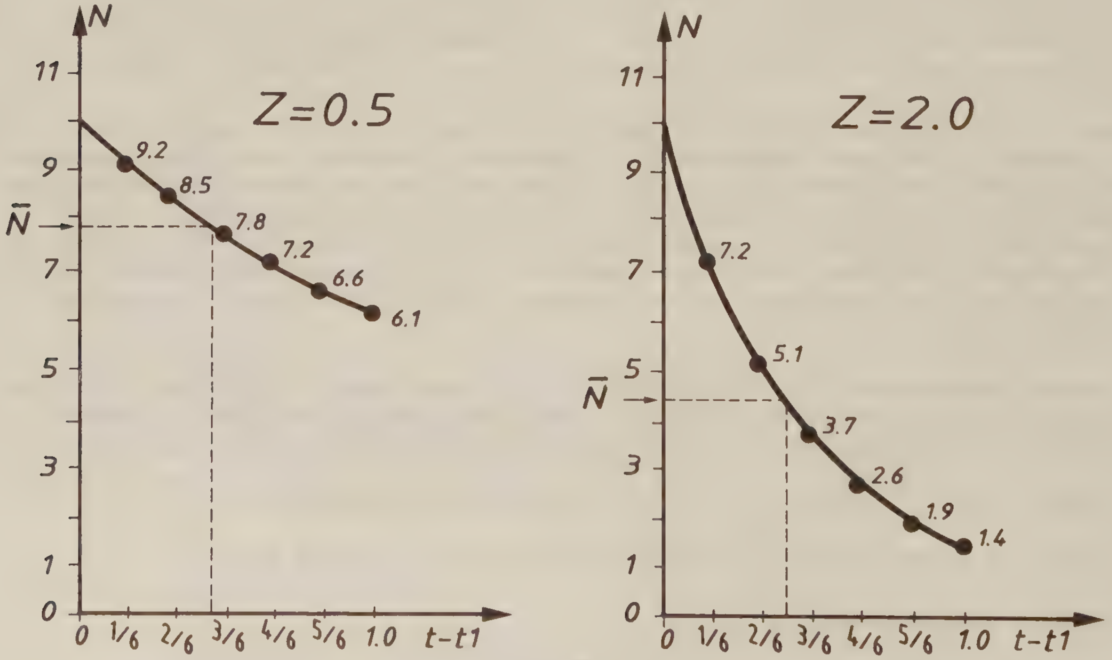
$$= 41269 \quad \text{(compare with Fig. 4.2.2)}$$

$$C(Tr+0.4,Tr+0.8) = (0.8-0.4) * 0.6 * 41269 = 9905,$$

which is also the answer we got from applying Eq. 4.2.4.

Eq. 4.2.8 has the advantage of being easy to interpret. It says that the number caught during the time period from t1 to t2 depends on:

1) The length of the time period
2) The fishing mortality
3) The average number of fish in the sea



Z = 0.5 Approximation: $\bar{N}$ = (10 + 9.2 + 8.5 + 7.8 + 7.2 + 6.6 + 6.1)/7 = 7.9
Exact expression: $\bar{N}$ = 10 * [1 - exp(-0.5)]/0.5 = 7.9
Z = 2.0 Approximation: $\bar{N}$ = (10 + 7.2 + 5.1 + 3.7 + 2.6 + 1.9 + 1.4)/7 = 4.5
Exact expression: $\bar{N}$ = 10 * [1 - exp(-2.0)]/2.0 = 4.3

Fig. 4.2.3 Illustration of Eq. 4.2.9:
N(t1,t2) = N(t1) [1 - exp(-Z(t2-t1))]/Z if t2-t1 = 1 year and Z = 0.5 or 2.0. Note that the larger the Z is the larger is the deviation between approximation and exact value.

Table 4.2.2 **Example of the dynamics of a cohort with different mortality rates during its lifespan**

start of period t1	end of period t2	natural mort. M	fishing mort. F	total mort. Z	comments
Tr	Tr+0.5	2.0	0	2.0	The cohort is not yet recruited to the fishing grounds. Exposed to great predation.
Tr+0.5	Tr+1.0	1.0	0.5	1.5	The cohort has migrated to fishing grounds, but 50% escape through the meshes. Predation mortality reduced.
Tr+1.0	Tr+7.0	0.5	1.0	1.5	The cohort under full exploitation (all fish are retained by the meshes). Predation mortality further reduced.

and each of these statements is easy to accept.

If /\t = t2-t1 is very small, it can be shown that approximately

$$C(t,t+/\backslash t) = /\backslash t\ F\ N(t) \qquad (4.2.10)$$

If /\t is small the number of survivors will change very little during the period and N and $\bar{N}$ will be approximately equal. Eq. 4.2.10 then follows from Eq. 4.2.8. Eq. 4.2.10 is a third version of the catch equation which is convenient for reasoning about the fishing mortality created by a single fishing operation, e.g. one trawl haul, as will be demonstrated in Section 4.3.

To apply the exponential model, Eq. 4.2.2 and the catch equations, it is not necessary to assume that M, F and Z remain constant during the entire life span of the cohort. The life span may be divided into shorter time periods within which mortalities are assumed to remain constant whereas they may vary form period to period. As an example, consider a cohort with a life span of seven years. The seven years may be divided into three periods as shown in Table 4.2.2. Suppose that N(Tr) = 100000 then, using Eq. 4.2.6:

$$\begin{aligned}
N(Tr+0.5) &= 100000 \exp(-2.0 * 0.5) = 36788\\
N(Tr+1.0) &= 36788 \exp(-1,5 * 0.5) = 17377\\
N(Tr+6.0) &= 17377 \exp(-1.5 * 5.0) = 9.61\\
N(Tr+7.0) &= 9.61 \exp(-1.5 * 1.0) = 2.14\\
N(Tr+8.0) &= 2.14 \exp(-1.5 * 1.0) = 0.48
\end{aligned}$$

Thus, after seven years the cohort has died out. The number caught in the first period (from Tr to Tr+0.5) is zero because F = 0. The number caught in the second period (from Tr+0.5 to Tr+1.0) is (using Eq. 4.2.7):

$$C(Tr+0.5,Tr+1) = N(Tr+0.5)\ \frac{F}{Z}\ (1 - \exp(-Z * 0.5))$$

$$= 36788\ \frac{0.5}{1.5}\ (1 - \exp(-1.5 * 0.5)) = 6470$$

Alternatively, Eq. 4.2.4 could have been used to calculate C:

$$C(Tr+0.5,Tr+1.0) = \frac{F}{Z}\,(N(Tr+0.5) - N(Tr+1.0)) =$$

$$\frac{0.5}{1.5}\,(36788 - 17377) = 6470$$

The third period is treated in a similar way. The results may be summarized as follows:

t1	t2	t2-t1	Z	N(t1)	F	C(t1,t2)
Tr	Tr+0.5	0.5	2.0	100000	0	0
Tr+0.5	Tr+1.0	0.5	1.5	36788	0.5	6470
Tr+1.0	Tr+7.0	6.0	1.5	17377	1.0	11583

(See **Exercise(s)** in Chapter 17).

4.3 Estimating Z from CPUE data and the concept of the catchability coefficent

The total mortality coefficient Z can be estimated when estimates of the number of fish in the sea are available for two different moments, t1 and t2. To calculate Z from such data Eq. 4.2.6 may be rewritten:

$$\frac{1}{t2-t1}\ln\frac{N(t1)}{N(t2)} = Z \tag{4.3.0.1}$$

However, it is not necessary to know the absolute value of N, only the ratio is required. Thus, if an estimate of some quantity proportional to N is available Eq. 4.3.0.1 can still be applied.

The catch per unit of effort (CPUE), for example the numbers caught per trawl hour, is a quantity which can be assumed proportional to the number of fish in the sea, N. Intuitively, it is not hard to accept that if there are twice as many fish in the sea, twice as many will be caught per fishing operation. Let CPUE(t) be the number caught per unit of effort. The first case to consider deals with CPUE data obtained from a research trawl survey of short duration (say, one week). Mathematically, the assumption is expressed:

$$CPUE(t) = q\,N(t) \tag{4.3.0.2}$$

where q is a parameter called the "catchability coefficient". The more efficient the gear is, the higher the value of q because q is a measure of the ability of the gear to catch fish. From Eq. 4.3.0.2 it follows that:

$$\frac{N(t1)}{N(t2)} = \frac{q\,N(t1)}{q\,N(t2)} = \frac{CPUE(t1)}{CPUE(t2)}$$

Inserting the above into Eq. 4.3.0.1 gives:

$$\frac{1}{t2-t1}\ln\frac{CPUE(t1)}{CPUE(t2)} = Z \tag{4.3.0.3}$$

Eq. 4.3.0.3 can be used in time series of CPUE data from research surveys if age compositions have been obtained, either from reading ring structures in hard parts or from length frequency analysis. Recall that the number in each cohort can be estimated for instance by the Bhattacharya method (Section 3). Just as one can follow the progression of modes for the estimation of growth parameters (cf. Section 3.4.2) so can the total mortality be estimated when the frequencies are in the form of numbers caught per unit of effort. A special application of this approach is illustrated in Exercise 4.3.1.

Sometimes CPUE is derived as the average value over a longer period. This is the case when data are collected from a commercial fishery, where data are most often given on a quarterly or yearly basis. The model to be used is mathematically equivalent to Eq. 4.3.0.3, but conceptually it is slightly different. The parallel to Eq. 4.3.0.2 corresponding to a longer time period is:

$$\overline{CPUE}(t1,t2) = q\ \bar{N}(t1,t2) \qquad (4.3.0.4)$$

This is the model relevant to data on commercial fisheries. $\overline{CPUE}$ is usually calculated as the catch during the period from t1 to t2 divided by the effort for the period. $\bar{N}$ is the average number of survivors during the period from t1 to t2.

The theory behind Eqs. 4.3.0.2 and 4.3.0.4 will be used several times in the following chapters. We therefore use the rest of this section for a discussion of the reasoning leading to these expressions. Let f be the effort during one year (t2-t1 = 1.0). Eq. 4.3.0.4 may then be written:

$$C/f = q\ \bar{N} \qquad (4.3.0.5)$$

When t2-t1 = 1 year we have a special version of Eq. 4.2.8:

$$C = F\ \bar{N} \qquad (4.3.0.6)$$

Inserting Eq. 4.3.0.6 into Eq. 4.3.0.5 gives

$$F = q\ f \qquad (4.3.0.7)$$

Eq. 4.3.0.7 can be shown to apply to any time period [t2,t1]. It is intuitively easy to accept. It says that the more effort (e.g. boat days) goes into the fishery, the higher will be the fishing mortality or: the fishing mortality is proportional to the effort. As Eq. 4.3.0.7 is so simple we could have taken that as the starting point for the derivation of Eq. 4.3.0.2 or Eq. 4.3.0.4.

Fig. 4.3.0.1 illustrates Eq. 4.3.0.7 in a simplified system, consisting of 2 million fish uniformly distributed over a coastal area. The figure shows only 200 fish(1/10000 of the system).

The rectangle labeled "A" indicates the area swept by the trawl in, say, one hour. The catch is six fish. For simplicity, we assume that no fish dies due to natural causes during the towing period and that all fish encountered by the trawl are retained. Thus, the fishing mortality, Δt F1, created by haul A must satisfy Eq. 4.2.10 (in this case Δt is small):

$$\Delta\ C1 = \Delta t\ F1\ N \quad \text{or} \qquad (4.3.0.8)$$

$$6 = \Delta t\ F1\ 2000000 \quad \text{so that}$$

$$\Delta t\ F1 = 0.000003$$

Similarly, for two and three tows during one hour we get:

two tows (A+B) : $\Delta t\ F2 = 0.000006$
three tows (A+B+C) : $\Delta t\ F3 = 0.000009$

Let one trawl hour be one effort unit. Then from Eq. 4.3.0.7 we have:

one tow : $\Delta t\ F1 = \Delta t\ q * 1$
two tows : $\Delta t\ F2 = \Delta t\ q * 2$
three tows : $\Delta t\ F3 = \Delta t\ q * 3$

and q thus remains constant for different F-values:

$$q = \frac{\Delta t\ F1}{1\Delta t} = \frac{\Delta t\ F2}{2\Delta t} = \frac{\Delta t\ F3}{3\Delta t} = \frac{0.000003}{1/(365*24)} = 0.002628$$

The above reasoning applies in the case where the fish are uniformly distributed over the fishing grounds. However, this is rarely the case. For schooling fish (pelagics) the argument is simply wrong. When shooting the net at random the catch of a schooling fish species might be zero in nine out of ten tows, yet one might yield a very high catch. For demersal fish species the reasoning may be approximately correct although demersal fish also tend to aggregate in certain areas, which may or may not be also the fishable areas. The reasoning illustrated by Fig. 4.3.0.1 is behind the "swept area method" described in Chapter 13. That the reasoning is based on a relatively small catch and very few effort units, however, does not limit its applicability.

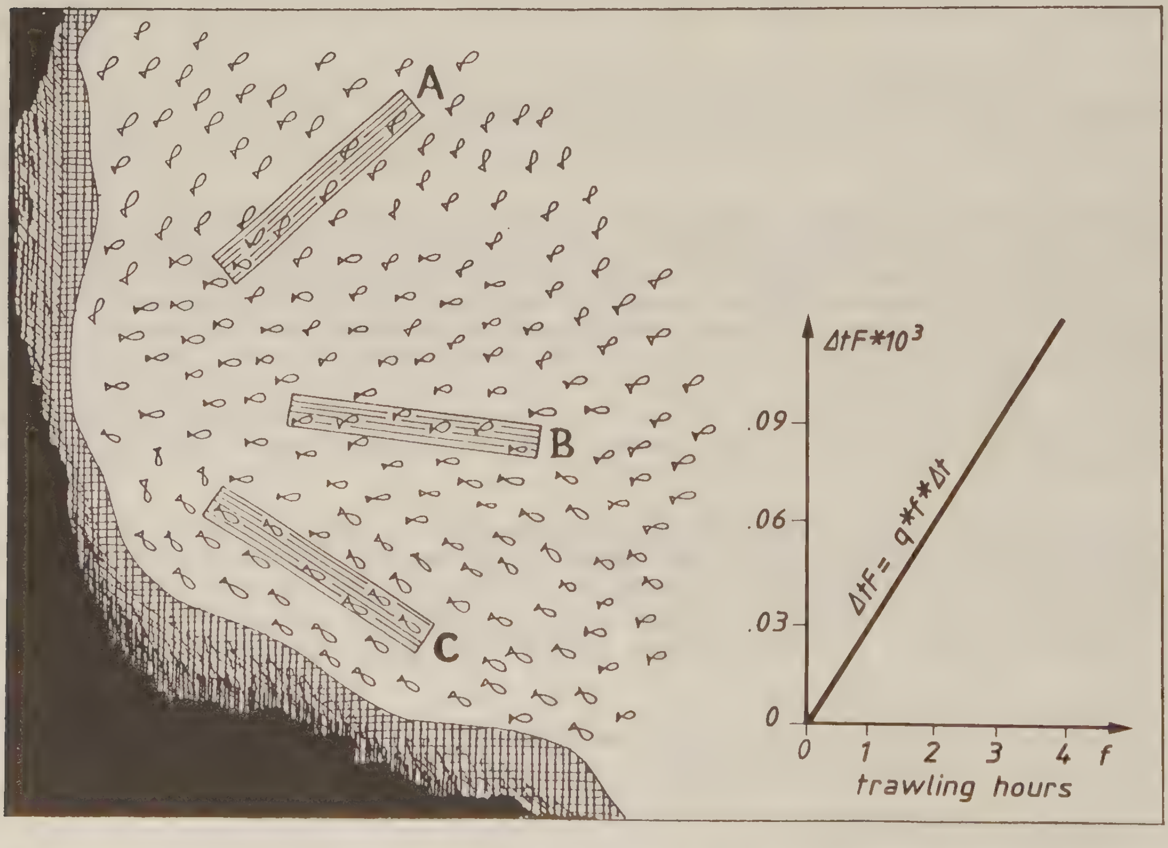


Fig. 4.3.0.1 Illustration of Eq. 4.3.0.7: F = q f. (For further explanation, see text)

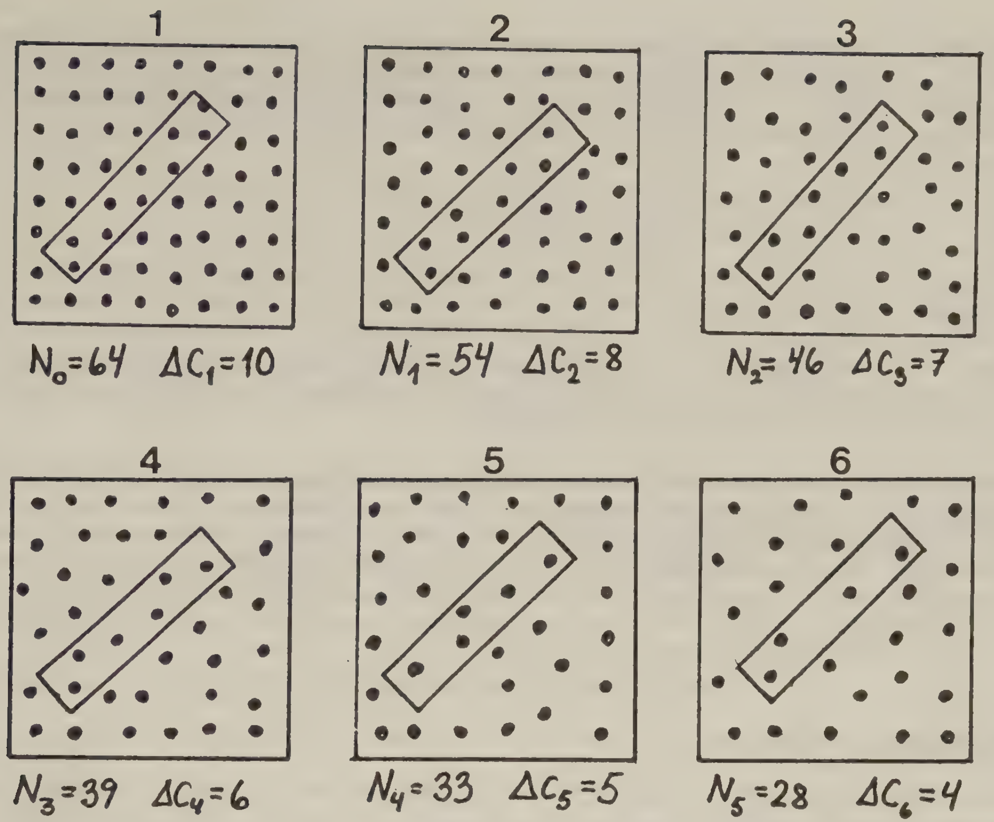

1
2
3
$N_0=64$ $\Delta C_1=10$
$N_1=54$ $\Delta C_2=8$
$N_2=46$ $\Delta C_3=7$
4
5
6
$N_3=39$ $\Delta C_4=6$
$N_4=33$ $\Delta C_5=5$
$N_5=28$ $\Delta C_6=4$

Fig. 4.3.0.2 Illustration of Eq. 4.3.0.7, $F = q f$, applied several times to a system during a longer time period. (For further explanation, see text)

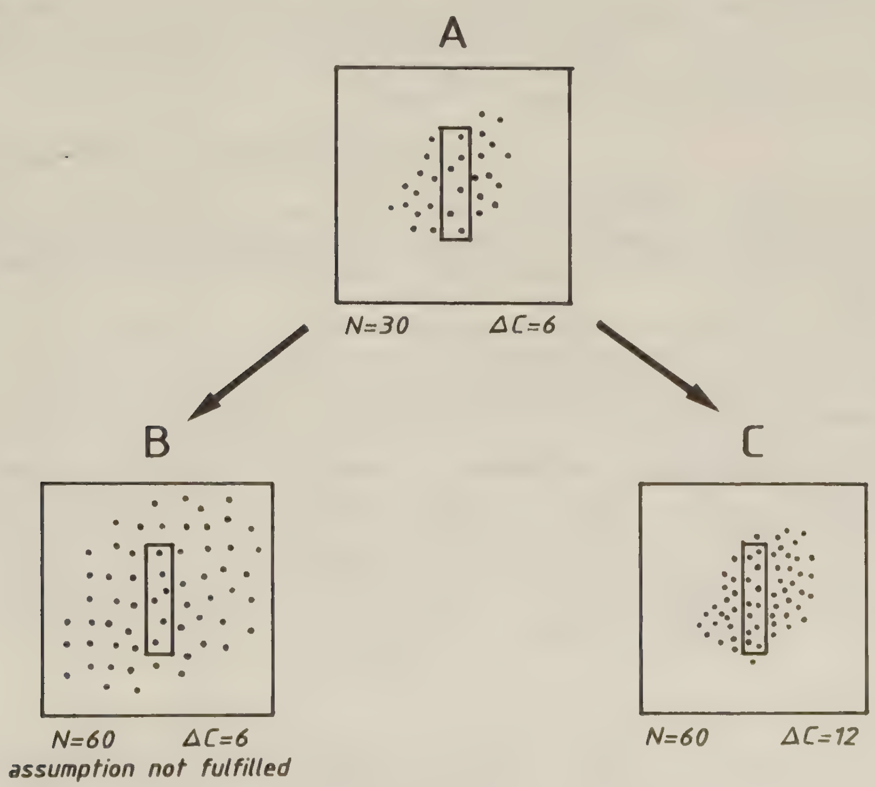

A
$N=30$ $\Delta C=6$
B
$N=60$ $\Delta C=6$
assumption not fulfilled
C
$N=60$ $\Delta C=12$

Fig. 4.3.0.3 Illustration of one of the assumptions behind Eq. 4.3.0.7: $F = q f$. (For further explanation, see text)

Now consider a system during a longer period. Fig. 4.3.0.2 shows such a (hypothetical) example, where more than half of the fish are removed by fishing. Note that one unit of effort (one trawl haul) produces the same fishing mortality independent of the number of survivors, i.e. all hauls give equal $\Delta t\ F$:

$$\Delta t\ F \approx \frac{8}{54} \approx \frac{7}{46} \approx \frac{6}{39} \approx \frac{5}{33} \approx \frac{4}{28} \approx 0.15$$

where ≈ means "approximately equal to". One unit of effort (one trawl haul or tow) takes the same fraction from the stock. The fishing mortality is independent of the number of survivors, i.e., all hauls give equal $\Delta t\ F$. Note also that $\Delta t\ F$ will remain constant only if the fish redistribute uniformly after each removal by fishing.

Figs. 4.3.0.3A and B show an example in which Eq. 4.3.0.7, $F = q\ f$, is not applicable, viz. when an increase in numbers produces also an increase of the area occupied by the stock. Then one effort unit (one trawl haul) would remove a smaller fraction of the stock and thus produce a smaller fishing mortality. Figs. 4.3.0.3A and C illustrate the way a stock must behave to conform to Eq. 4.3.0.7, namely that the area occupied remains constant for different stock numbers.

4.3.1 Heincke's method

There are many ways of estimating Z from numbers in the catch. We shall describe only one that dates as far back as 1913. Several other methods are described in Ricker (1975).

To introduce the method we start with a formula from common algebra:

$$\text{For } a < 1 \text{ we have } \sum_{i=0}^{\infty} a^i = \frac{1}{1 - a}$$

$a = \exp(-Z)$ gives $a < 1$ for $Z > 0$ and

$$\sum_{t=0}^{\infty} \exp(-Zt) = \frac{1}{1 - \exp(-Z)} \qquad (4.3.1.1)$$

which is equivalent to (we multiply by the same constant on both sides of the equal sign):

$$\sum_{t=0}^{\infty} N(0) \exp(-Zt) = \frac{N(0)}{1 - \exp(-Z)} \qquad (4.3.1.2)$$

Assuming total mortality, Z, to remain constant we have:

$N(t) = N(0) \exp(-Zt)$ and

$$\sum_{t=0}^{\infty} N(t) = \frac{N(0)}{1 - \exp(-Z)} \qquad (4.3.1.3)$$

Equation 4.3.1.3 is equivalent to:

$$Z = -\ln \frac{N(1) + N(2) + N(3) + N(4) + \ldots\ldots\ldots}{N(0) + N(1) + N(2) + N(3) + N(4) + \ldots\ldots\ldots}$$

or

$$Z = -\ln\left[\sum_{t=1}^{\infty} N(t) \Big/ \sum_{t=0}^{\infty} N(t)\right] \qquad (4.3.1.4)$$

Eq. 4.3.1.4 is "Heincke's formula" (Heincke, 1913, in Ricker, 1975).

Replacing the N's by CPUE(t), Eq. 4.3.1.4 should still hold if CPUE can be assumed proportional to the stock number. Eq. 4.3.1.4 is usually used in the form:

$$Z = -\ln \frac{\text{CPUE}(1) + \text{CPUE}(2) + \text{CPUE(age 3 and older)}}{\text{CPUE}(0) + \text{CPUE}(1) + \text{CPUE}(2) + \text{CPUE(age 3 and older)}}$$

as it is often so that age groups can be relatively easily separated for the first 2 or 3 (or perhaps 4) age groups, whereas it is difficult for the older age groups (Ricker, 1975). It is sometimes necessary to exclude the youngest age group as being not fully exploited by the fishery, i.e., to leave out CPUE(0) in the denominator and CPUE(1) in the numerator.

4.4 Estimation of Z from a linearized catch curve

A catch curve is a graphical representation of numbers of survivors plotted against age.

The methods to be described in this section are essentially based on age composition data and as such they have been applied to temperate fish stocks. Because direct age reading is problematic for tropical fish stocks (cf. Section 3.2.1) the methods are here converted into versions which take length composition data as input. This is possible when growth parameters are available so that lengths can be converted into ages by using the inverse von Bertalanffy equation (Eq. 3.3.3.3). The theoretical derivations start by using the age structured model, which will subsequently be converted into a corresponding length structured model.

4.4.1 The constant parameter system

The linearized catch curve methods for estimating Z are based on the assumption of a "constant parameter system". This section is therefore used to explain this concept.

Consider, as an example, a fish stock during the period 1971-1975 with a life span of five years. Let the numbers of survivors be those given in Table 4.4.1.1A. For simplicity we assume only one cohort per year, recruiting on 1 January. The figures in Table 4.4.1.1A are the numbers of survivors per cohort and per age group, on 1 January. Because we consider many cohorts simultaneously, an index to specify the cohort, y, is needed. N(y,t) is the number of survivors attaining age t from the cohort recruited at the start of year y-t. In this particular case t only takes the values t = 0,1,2,3 or 4. Note that a specific cohort can be followed diagonally across the table. Every year the survivors of that cohort advance to a new age group. Each column, on the other hand, contains the numbers of survivors of five different cohorts 1 January of a particular year.

Table 4.4.1.1A contains parts of nine different cohorts, of which only one is given for its entire life span (the 1971-cohort). It appears that recruitment, N(y,0), y = 1971, 1972, 1973, 1974 and 1975, has varied from year to year. The 1970-cohort and the 1973-cohort are strong ones, whereas the 1972-cohort is a weak one ("strong" means that N(y,0) is considerably greater than the average recruitment).

Table 4.4.1.1 Illustration of the "constant parameter system" and the "variable parameter system". Note that the column of numbers in boxes consists of five different cohorts in the year 1971, whereas the diagonal of numbers in boxes consists entirely of the 1971-cohort. (For further explanation, see text)

A: VARIABLE PARAMETER SYSTEM

age\year t	1971 N(71,t)	1972 N(72,t)	1973 N(73,t)	1974 N(74,t)	1975 N(75,t)
0	$\boxed{2105}$	1111	9560	1869	1236
1	$\boxed{2575}$	$\boxed{736}$	405	3817	618
2	$\boxed{155}$	1097	$\boxed{281}$	142	1193
3	$\boxed{102}$	58	298	$\boxed{109}$	59
4	$\boxed{6}$	38	18	138	$\boxed{29}$

B: CONSTANT PARAMETER SYSTEM

age\year t	1971 N(71,t)	1972 N(72,t)	1973 N(73,t)	1974 N(74,t)	1975 N(75,t)
0	$\boxed{2560}$	2560	2560	2560	2560
1	$\boxed{942}$	$\boxed{942}$	942	942	942
2	$\boxed{346}$	346	$\boxed{346}$	346	346
3	$\boxed{127}$	127	127	$\boxed{127}$	127
4	$\boxed{47}$	47	47	47	$\boxed{47}$

Now, suppose that the recruitment, N(y,0), remains constant every year, as is the case in Table 4.4.1.1B. Assume further that also F and M remain constant, then the number of survivors and the numbers caught would be the same for all cohorts. An inspection of Table B shows that in this case the number of survivors per year during the life span of a cohort equals the number of survivors within a particular year for each age group. Thus, in the case of constant recruitment it does not matter whether we consider one cohort over its entire lifespan, or if we consider all the different cohorts (the entire stock) in one particular year. The major part of the Beverton and Holt theory (1957) is based on the assumption of constant recruitment.

We know for sure that the assumption of a constant parameter system is never strictly fulfilled in real life. However, we are often in a situation which forces us to make assumptions, which are known to be crude approximations to reality. It often happens that only by making such assumptions we are able to carry out an analysis of available data, and it is better to do a crude analysis than to do no analysis at all.

We speak of a "pseudo-cohort" when working with data from one year assuming these to resemble those of a cohort during its life-span. Thus, the numbers 2105, 2575, 155, 102 and 6 in Table 4.4.1.1A form a pseudo-cohort, whereas the numbers 2105, 736, 281, 109 and 29 form a real cohort.

4.4.2 The linearized catch equation

Suppose that an estimate of the age composition of the catch during a year is available, i.e. pairs of observations:

age interval	numbers caught
t1 - t2	C(t1,t2)
t2 - t3	C(t2,t3)
t3 - t4	C(t3,t4)
.	.
.	.

To develop a method to estimate Z from such data, the starting point is the catch equation (Eq. 4.2.7):

$$C(t,t2) = N(t1)\,\frac{F}{Z}\,[1 - \exp(-Z(t2-t1))]$$

Although the time interval from t1 to t2 is the first interval it will also be used in the following as the symbol for the general time interval, i.e. t2-t3 or t3-t4 or .... etc. Eq. 4.2.7 is not linear in t1 (or t2), so some transformation is required to make the analysis a linear regression (cf. Section 2.6). The first step is to use Eq. 4.2.2 which gives

$$C(t1,t2) = N(Tr)\,\exp(-Z(t1-Tr))\,\frac{F}{Z}\,[1 - \exp(-Z(t2-t1))]$$

or by an ln-transformation:

$$\ln C(t1,t2) = \ln\left(N(Tr)\,\frac{F}{Z}\right) + ZTr - Zt1 + \ln[1 - \exp(-Z(t2-t1))]$$

At least t1 now appears in one linear term (namely: "-Z t1"). The term "ln(N(Tr)F/Z)+ZTr" is a constant, as N(Tr), Tr, F and Z are assumed to remain constant. To simplify the notation, this constant is named "d", and the equation reads:

Table 4.4.3.1 Linearized catch curve analysis based on age composition data. Number caught per year by age group (in millions caught per year) of North Sea whiting. (From ICES, 1980, compare Fig. 4.4.3.1)

year y age t	1974	1975	1976	1977	1978	1979	1980	average 1974-80	
				C(y,t,t+1)					
0	599	239	424	664	685	478	330	488	not
1	678	860	431	1004	418	607	288	612	used
2	1097	390	1071	532	335	464	323	601	
3	275	298	159	269	203	211	243	237	used
4	40	54	75	32	69	86	80	62	in
5	6	9	13	18	8	25	31	15.7	ana-
6	1	8	3	5	5	3	8	4.7	ly-
7	(6)	(0)	(1)	(0)	(1)	(1)	(1)	1.4	sis
Z	1.88	1.26	1.36	1.25	1.33	1.40	1.12	1.28	

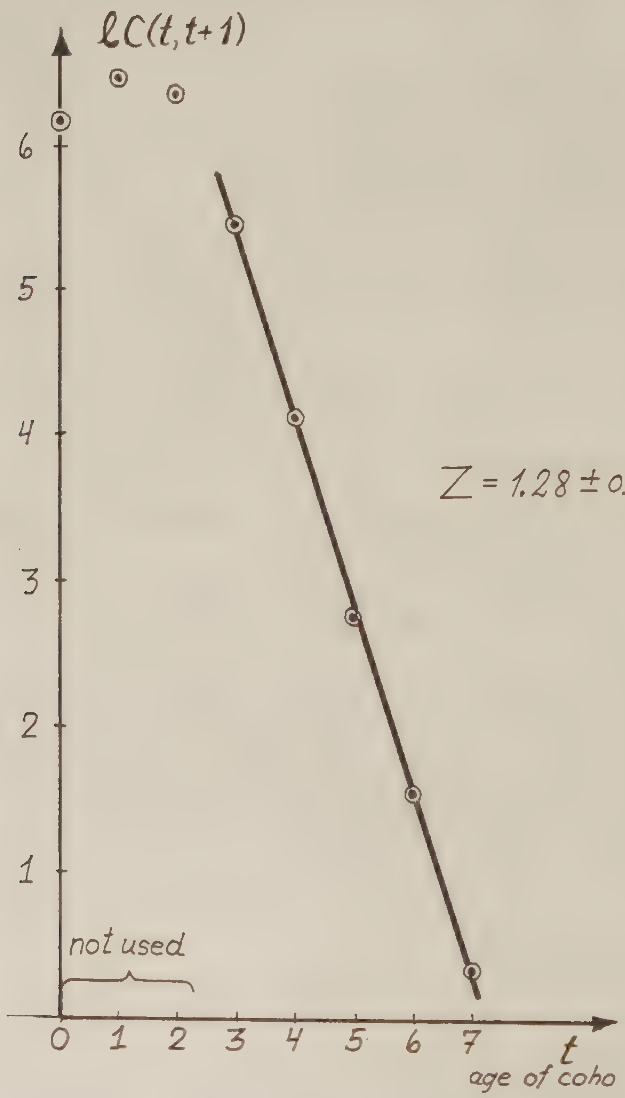


Fig. 4.4.3.1 Linearized catch curve based on average age composition of North Sea whiting catches 1974-80 (ICES, 1980)

$$\ln C(t1,t2) = d - Z\,t1 + \ln [1 - \exp(-Z(t2-t1))] \qquad (4.4.2.1)$$

Except for the last term a linear expression in t1 has been obtained. Various methods suggest different ways of dealing with this term.

The first uses of the linearized catch curve date back to Edser (1908), Heincke (1913) and Baranov (1918) as reviewed in Ricker (1975). In the remaining part of Section 4.4 we shall discuss various special applications of Eq. 4.4.2.1.

4.4.3 The linearized catch curve based on age composition data

One of the methods commonly applied in temperate waters to estimate total mortality is the "linearized catch curve method (with constant time intervals)", which has been reviewed in Beverton and Holt (1956), Chapman and Robson (1960), Robson and Chapman (1961) and Ricker (1975).

If t2-t1 in Eq. 4.4.2.1 remains constant (e.g. when age groups are considered, so that t2-t1 = t3-t2 = t4-t3 = ... = 1 year), the non-linear term becomes a constant, which can be included in the intercept term. Thus, if the constant g is defined:

$$g = d + \ln (1 - \exp(-Z(t2-t1)))$$

then Eq. 4.4.2.1 reads:

$$\ln C(t1,t2) = g - Zt \quad \text{or}$$

$$\ln(C(t,t+\Delta t)) = g - Z\,t \qquad (4.4.3.1)$$

Eq. 4.4.3.1 is the "linearized catch curve equation (with constant time intervals".

Table 4.4.3.1 shows an example of a linearized catch curve analysis based on age-group composition data. Most often such examples come from temperate waters where direct age reading is possible and where spawning is confined to a short period once per year. This is also the case here, where the example deals with whiting (*Merlangius merlangus*) caught in the North Sea.

Table 4.4.3.1 gives the annual age composition of whiting catches during the period 1974-1980 (from ICES, 1980). The figures in Table 4.4.3.1 are:

C(y,t,t+1) = number caught in year y of age between t and t+1 years (in millions). For instance C(1976,3,4) = 159

For each year the total mortality, Z, has been calculated using Eq. 4.4.3.1 (bottom row of Table 4.4.3.1). Thus, recruitment and other parameters are assumed to remain constant from 1974 to 1980.

However, we could also calculate an overall estimate of Z for the seven year period by using the average number caught per age group, as shown in the last column of Table 4.4.3.1. By doing so we (more or less) circumvent the problem of assuming parameters to remain constant, as e.g. variation in recruitment is leveled out in the averaging process. Note that in this particular case the assumption of constant recruitment is not crucial. Except for the year 1974 the annual estimates of Z do not deviate much from the overall estimate.

Fig. 4.4.3.1 shows the graph for the estimation of the overall Z. As appears from Table 4.4.3.1 data for the three first age groups have been excluded from the regression analysis. From Fig. 4.4.3.1 it is seen that they do not conform to the straight line. The reason is that these age groups are not under full exploitation. The fishing mortality is lower than that of the larger (older) fish because age groups 0-2 are so small that they escape through the meshes of the trawl. (The major part of the North Sea whiting are caught by trawls with codend mesh sizes from 70 mm to 110 mm). Another reason for expecting a reduced fishing mortality on the small fish is that all of these may not yet have recruited to the fishing grounds and are still on the nursery grounds.

Thus total mortality, $Z = M+F$, does not remain constant for all age groups although this was the assumption behind the linearized catch curve analysis. One way to circumvent the problem is to exclude those age groups we suspect not to be under full exploitation.

Which age groups to exclude is a somewhat subjective choice. Usually the points systematically deviating from the straight line are excluded. However, it is hard to give a general rule for when this deviation is sufficiently large to justify the exclusion of the point. As a rule of thumb: whenever you are in doubt about a point, exclude it. In the case of Fig. 4.4.3.1 it is easy to make the choice since the five points used provide a very convincing fit to the model.

In the estimation of Z from one year's data (bottom row of Table 4.4.3.1) age group 7 has been excluded as well (indicated by brackets). In this case the reason is that the number of 7+ groups are estimated from small samples (simply because so few fish survive to age 7) and therefore they are subject to great uncertainty. For the overall estimate of Z, age group 7 has been used, because in this case the samples extend to seven years and the sample size is accordingly larger.

Using the procedure described in Section 2.4 the 95% confidence limits of Z can be calculated:

$$Z = -\text{ slope} = 1.284 \quad sy = 2.0316 \quad sx = 1.5811$$

variance of the slope:

$$sb^2 = \frac{1}{n-2}\left((sy/sx)^2 - b^2\right) =$$

$$= \frac{1}{5-2}\left((2.031576/1.581139)^2 - 1.28430^2\right) = 0.0004979$$

$sb\ t(5-2) = 0.0223 * 3.18 = 0.07$, so that the confidence interval becomes

$$[1.21, 1.35] \quad \text{or: } Z = 1.28 \pm 0.07$$

With the data of Table 4.4.3.1 the calculations could have been made on a cohort basis as well. Exercise 4.4.3 deals with this aspect.

(See **Exercise(s)** in Chapter 17).

4.4.4 The linearized catch curve based on age compositions with variable time intervals

Let us return to the general linearized catch curve model (Eq. 4.4.2.1)

$$\ln C(t1,t2) = d - Z\ t1 + \ln\left[1 - \exp(-Z(t2-t1))\right]$$

In case the time intervals t2-t1, t3-t2, t4-t3, ... are not constant the non-linear term, $\ln [1 - \exp(-Z(t2-t1))]$ takes different values for different values of the time interval. As will appear from Section 4.4.5 this case is often relevant to tropical fish stocks. Two alternative ways of getting rid of the non-linear term have been suggested.

Jones and van Zalinge (1981) suggested to let t2 take a very high value (actually $t2 = \infty$ (infinity)). With t2 very large, the term $\exp(-Z(t2-t1))$ would be close to zero, and consequently $\ln(1 - \exp(-Z(t2-t1))) = \ln 1 = 0$. Thus, if $C(t,\infty)$ designates all fish caught of age t and older, a linear relationship is achieved by

$$\ln C(t,\infty) = d - Z\ t \qquad (4.4.4.1)$$

$C(t,\infty)$ is called the "cumulated catch", and Eq. 4.4.4.1 the "cumulated catch curve equation".

Van Sickle (1977) suggested a different approach which was further developed by Pauly (1983a). It is based on the observation that for small values of x, ($x < 1.0$):

$$\ln(1 - \exp(-x)) = \ln x - x/2 \quad \text{(approximately)}$$

If $Z(t2-t1)$ is small it follows that approximately:

$$\ln[1 - \exp(-Z(t2-t1))] = \ln[Z(t2-t1)] - Z(t2-t1)/2$$

Inserting this into Eq. 4.4.2.1 and rearranging terms gives:

$$\ln \frac{C(t1-t2)}{t2-t1} = d + \ln Z - Z\ t1 - Z(t2-t1)/2$$

Because t2-t1 must be small, it is called $\underline{\Delta}t$. To simplify notation the constant $c = d + \ln Z$ is introduced. Then the equation reads:

$$\ln \left[\frac{C(t,t+\underline{\Delta}t)}{\underline{\Delta}t} \right] = c - Z\ (t + \underline{\Delta}t/2) \qquad (4.4.4.2)$$

which is linear in the interval midpoint, $(t + \underline{\Delta}t/2)$. Eq. 4.4.4.2 is the so-called "linearized catch curve equation (with variable time intervals)". Note that only the slope has significance to the present analysis. The intercept is not used. The name of the equation might lead to the misunderstanding that the method is applicable only when data for total catches are available. However, Eq. 4.4.4.2 can still be used even if only the percentage composition of the catches is known, because from these we get the absolute catches by multiplying all the percentages by a constant.

A constant, K, multiplied to all the values for C in, Eq. 4.4.4.2 will change the intercept, c, but not the slope, Z, since

$$\ln\ (K * C/\underline{\Delta}t) = c - Z\ (t+\underline{\Delta}t/2) \quad \text{is equivalent to:}$$

$$\ln\ (C/\underline{\Delta}t) = c1 - Z\ (t+\underline{\Delta}t/2) \quad \text{where}$$

$$c1 = c - \ln K$$

Thus, if you have a sample from a fish population (for instance from a research survey or a market), but the total catch is unknown, you can use Eq. 4.4.4.2. Therefore the linearized catch curve analysis is often used for survey data, whereas cohort analysis (introduced in Chapter 5) is used for the analysis of total catch data. It should be noted that the example used to illustrate a linearized catch curve analysis (cf. Table 4.4.5.1) is a sample of only 3816 fish, which was not raised to the total catch.

4.4.5 The linearized catch curve based on length composition data

The method to be described here does not as Eq. 4.4.3.1, assume a direct age reading but uses the von Bertalanffy growth equation to convert length into age. This model is discussed in Pauly (1983a, 1984a and b). It is often referred to as the "length converted catch curve" or the "linearized length converted catch curve".

To make the catch curve usable for length data it is in principle necessary to replace the t's (ages) in Eq. 4.4.4.2

$$\ln \frac{C(t,t+\Delta t)}{\Delta t} = c - Z\,(t+\Delta t/2)$$

by L's (lengths).

However, what is actually done is to convert length data into age data, using the inverse von Bertalanffy growth equation (Eq. 3.3.3.3)

$$t(L) = t_o - \frac{1}{K}\ln\left(1 - \frac{L}{L_\infty}\right)$$

The conversion of lengths into ages is fairly complex, because the amount of time needed for a fish to grow through a given length interval increases continuously as illustrated in Fig. 4.4.5.1.

The age corresponding to a certain length can be calculated, when t_o is ignored or assumed to be 0.

Let age t correspond to L1 and age $t+\Delta t$ to L2:

$$L1 = L(t) \text{ and } L2 = L(t+\Delta t)$$

With this one to one correspondence between the age interval $[t,t+\Delta t]$ and the length interval [L1,L2] we can change the notation for the number caught:

$$C\,(t,t+\Delta t) = C\,(L1,L2)$$

In the present context Δt is the time it takes for an average fish to grow from length L1 to length L2, so we obtain Δt by subtracting the two inverse von Bertalanffy equations (Eq. 3.3.3.3) corresponding to L2 and L1 respectively and obtain:

$$\Delta t = t(L2) - t(L1) = \frac{1}{K}\ln\frac{L_\infty - L1}{L_\infty - L2} \qquad (4.4.5.1)$$

It is now possible to calculate Δt directly from the growth parameters K, and L_∞ and the corresponding L1 and L2, as well as, of course by subtracting t(L1) from t(L2).

The term $(t+\Delta t/2)$ of Eq. 4.4.4.2 can also be converted into an expression in length (L1 and L2), based on a suggestion by Pauly, viz., that age interval $(t+\Delta t/2)$ midpoint can be converted into a length based midpoint by assuming that

$t(L1) + \Delta t/2$ is approximately equal to

$$\frac{t(L1) + t(L2)}{2} = t_o - \frac{1}{K}\ln\left(1 - \frac{L1+L2}{2L_\infty}\right) \qquad (4.4.5.2)$$

The final result is the so-called "(linearized) length converted catch curve":

$$\ln \frac{C(L1,L2)}{\Delta t(L1,L2)} = c - Z \left(\frac{t(L1) + t(L2)}{2}\right) \qquad (4.4.5.3)$$

This is a linear equation where

$$Y = \ln \frac{C(L1,L2)}{\Delta t(L1,L2)} \quad \text{and} \quad X = \frac{t(L1 + t(L2)}{2}$$

while the slope (b) is -Z.

Thus, given a set of length frequency data and the growth parameters K and L_∞ it is possible to obtain an estimate of Z.

An example of its use is given in Table 4.4.5.1 based on data for *Upeneus vittatus* from Manila Bay (Ziegler, 1979). The input data are columns A and B, as well as values for L_∞ and K. Then column C, t(L1), is calculated by Eq. 3.3.3.3 and column D either from column C or by Eq. 4.4.5.1, while X (column E) is derived from column C, and Y (column F) from columns B and D.

The next step is to plot X against Y (Fig. 4.4.5.2) and to decide which points should be used for a regression analysis, of which the slope b will yield -Z.

The confidence limits are calculated for various numbers of observations, so as to determine the best estimate of Z. (A similar procedure should be followed in Exercise 4.4.5).

As was the case for the age-group based linearized catch curve, some of the observations have been excluded from the regression analysis in Table 4.4.5.1. The first seven length groups (6 - 13 cm) form the ascending part of the curve. These fish are considered not yet fully recruited to the fishery. The last four length groups are excluded as well, the reasons being:

1) Small numbers in the samples (as for the age-group based linearized catch curve)

2) When approaching L_∞, the relationship between age, t(L), and length, L, becomes uncertain (cf. the problems with the von Bertalanffy plot discussed in Section 3.3.3).

Point two above is an important one for length based linearized catch curves. It is a good reason for never using the largest length groups.

Fig. 4.4.5.1 shows the transformation of length groups into age intervals. The growth curve used is the average growth curve for the entire cohort, from which large individual deviations can occur. These deviations are relatively small for small fish, but may be relatively large for large fish. In other words: the larger the fish, the longer time it will take to grow through a length group and therefore the relationship between length and age becomes more inaccurate for large fish. The biggest fish in the catch are often bigger because they grow faster, not because they are older as we assume in the inverse von Bertalanffy equation. As an example consider the resolution of the length frequency sample in Fig. 3.2.2.2 into normally distributed cohorts. The length group 15-16 cm, consists only of members of the 1983 spring cohort (as appears from Table 3.2.1.1), whereas the length group 45-46 cm contains three cohorts:

1. Fast growing members of the 1981 autumn cohort
2. Medium/fast growing members of the 1981 spring cohort
3. Medium growing members of the 1980 autumn cohort

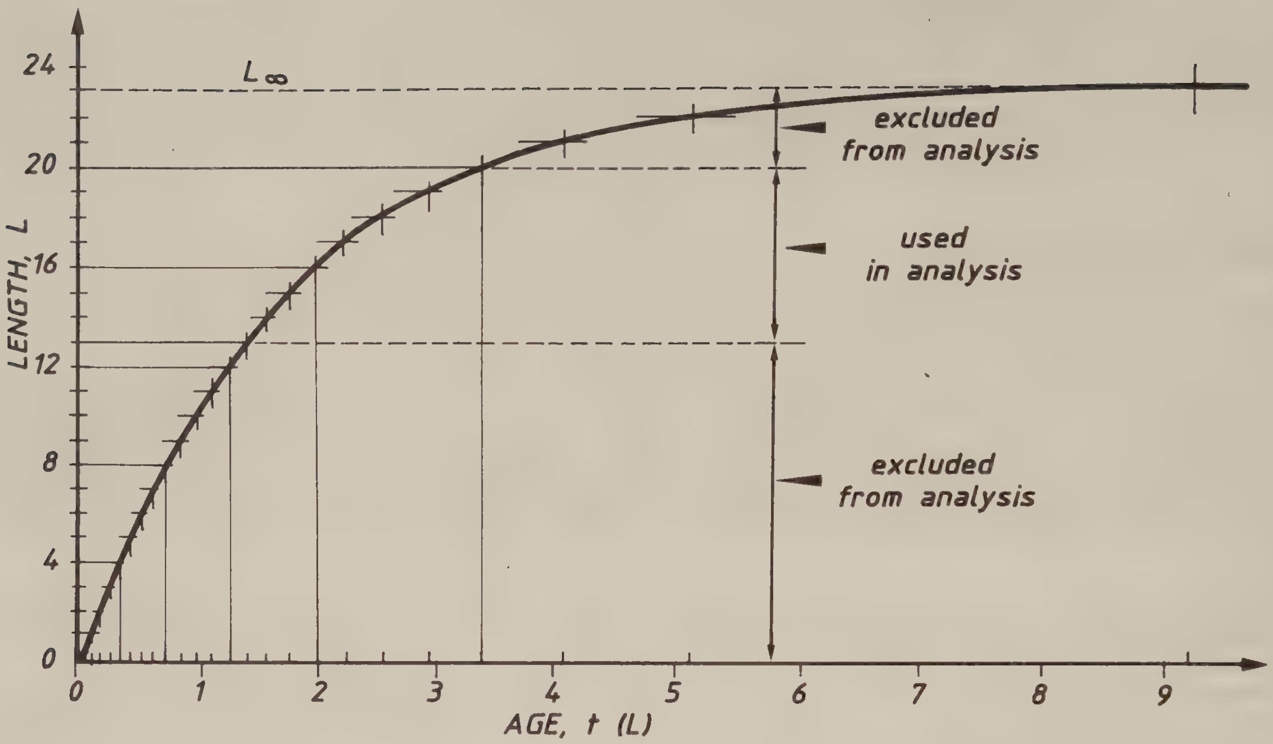


Fig. 4.4.5.1 Conversion of 1-cm length groups of *Upeneus vittatus* into age intervals, by the equation:

$$t(L) = t_o - \frac{1}{K}\ln(1 - L/L_\infty) \quad (t_o = 0,\ K = 0.59 \text{ and } L_\infty = 23.1)$$

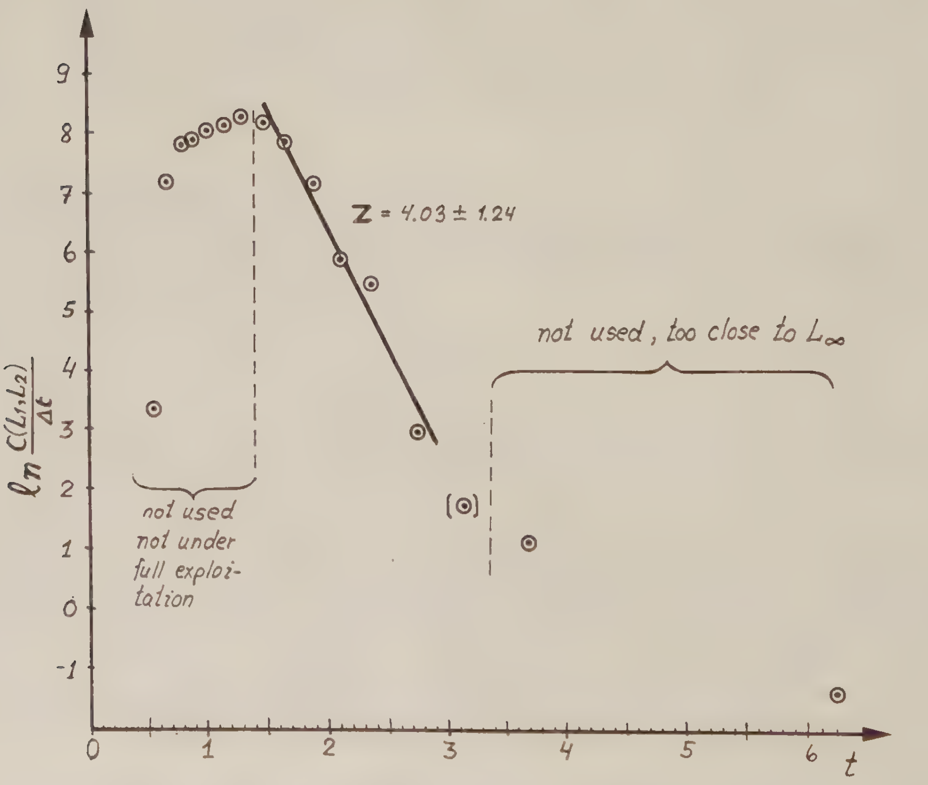


Fig. 4.4.5.2 Linearized catch curve based on length composition data for the yellow striped goatfish, *Upeneus vittatus* (from Ziegler, 1979). Compare Table 4.4.5.1

Table 4.4.5.1 **Linearized catch curve based on length composition data for the yellow striped goatfish, Upeneus vittatus from Manila Bay, Philippines (from Ziegler, 1979). L_∞ = 23.1 cm, K = 0.59 per year. (Compare Fig. 4.4.5.1)**

L1-L2	C(L1,L2)	t(L1)	Δt	$t(\frac{L1+L2}{2})$ X	$\ln \frac{C(L1,L2)}{\Delta t}$ Y	Z	remarks
6- 7	3	0.510	0.102	0.56	3.38	-	not
7- 8	143	0.612	0.109	0.67	7.18	-	used;
8- 9	271	0.721	0.116	0.78	7.76	-	not
9-10	318	0.837	0.125	0.90	7.86	-	under
10-11	416	0.961	0.134	1.03	8.04	-	full
11-12	488	1.096	0.146	1.17	8.11	-	exploi-
12-13	614	1.242	0.160	1.32	8.25	-	tation
13-14	613	1.402	0.177	1.49	8.15	-	portion
14-15	493	1.579	0.197	1.67	7.83	-	used
15-16	278	1.776	0.223	1.88	7.13	2.64	in the
16-17	93	2.000	0.257	2.12	5.89	3.61	regres-
17-18	73	2.257	0.303	2.40	5.48	3.20	sion
18-19	7	2.560	0.370	2.74	2.94	4.03	analy-
19-20	2	2.930	0.473	3.15	1.44	4.19	sis
20-21	2	3.404	0.659	3.70	1.11	-	not used;
21-22	0	4.064	1.094	4.53	-	-	too close
22-23	1	5.160	4.047	6.19	-1.40	-	to L_∞
23-24	1	-	-	-	-	-	

Details of the five regression analyses:

L1-L2	slope b Z	number of obs. n	Student's distrib. 95% t(n-2)	variance of slope sb^2	confidence limits Z ± t(n-2) sb
13-14	-	1	-	-	-
14-15	-	2	-	-	-
15-16	2.64	3	12.70	0.198	2.64 ± 5.65
16-17	3.61	4	4.30	0.283	3.61 ± 2.28
17-18	3.20	5	3.18	0.121	3.20 ± 1.11
18-19	4.03	6	2.78	0.198	4.03 ± 1.24
19-20	4.19	7	2.57	0.087	4.19 ± 0.76

Let us go back to the example of Table 4.4.5.1. Only the length groups 13-20 cm are used in the analysis of the data. To assess the sensitivity due to the choice of points to be excluded, the slope b = Z has been calculated from 3 points, 4 points, ..., and 7 points as shown at the bottom of Table 4.4.5.1. The first Z is calculated for the lengths 13-16 cm, the next Z for lengths 13-17 cm, ..., etc. Confidence limits have been calculated using the method described in Section 2.4. In the present case the conclusion is that Z is somewhere between 4.03-1.24 = 2.8 and 4.03+1.24 = 5.2 per year, using the six intervals from 13 to 19 cm. For the length groups 16-20 cm, $Z \Delta t > 1$ so that the approximation behind the linearized catch curve ($\ln(1-\exp(-x)) = \ln x - x/2$, cf. Section 4.4.4) is a crude one.

(See **Exercise(s)** in Chapter 17).

4.4.6 The cumulated catch curve based on length composition data. (The Jones and van Zalinge method)

To convert the "cumulated catch curve" into a length based analysis the inverse von Bertalanffy equation (Eq. 3.3.3.3)

$$t(L) = t_o - \frac{1}{K} \ln\left(1 - \frac{L}{L_\infty}\right)$$

is inserted into the cumulated catch curve equation (Eq. 4.4.4.1):

$$\ln C(t,\infty) = d - Zt$$

which gives the following results:

$$\ln(C(L1,L_\infty)) = d - Z\left(t_o - \frac{1}{K} \ln\left(\frac{1 - L}{L_\infty}\right)\right)$$

which can be converted into

$$\ln(C(L1,L_\infty)) = d - Zt_o - \frac{Z}{K} \ln(L_\infty) + \frac{Z}{K} \ln(L_\infty - L1)$$

The first three terms are constants which can conveniently be renamed:

$$a = d - Z\, t_o - \frac{Z}{K} \ln L_\infty$$

The "Jones and van Zalinge equation" has thus been derived:

$$\ln(C(L,L_\infty)) = a + \frac{Z}{K} \ln(L_\infty - L) \qquad (4.4.6.1)$$

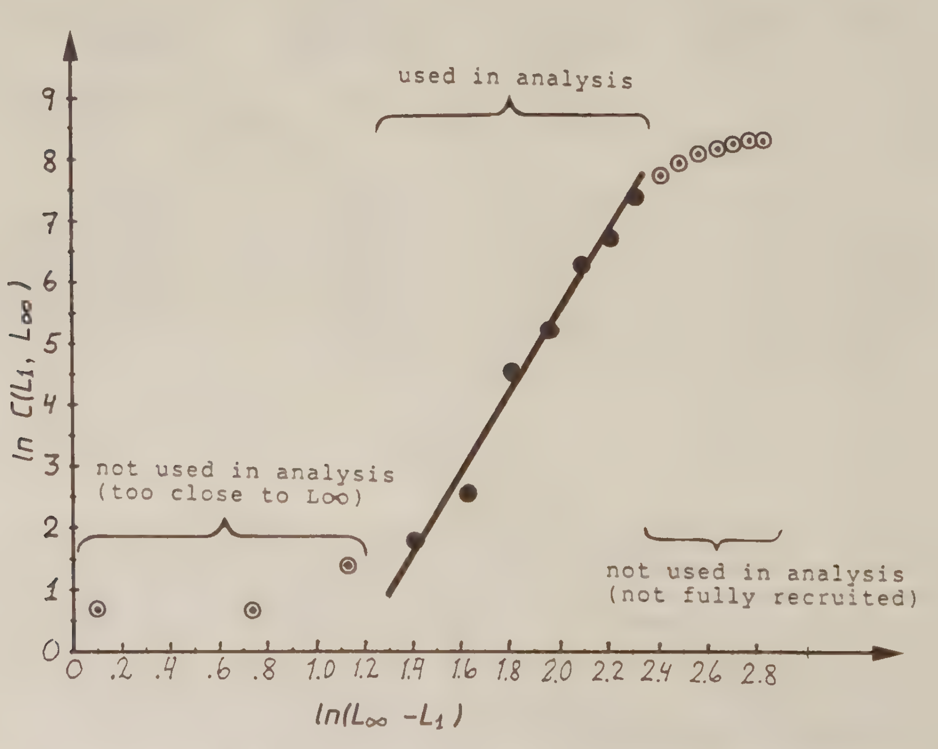


Fig. 4.4.6.1 Cumulated catch curve based on length composition data for the yellow striped goatfish, Upeneus vittatus. The line drawn is estimated from seven points. Data from Table 4.4.6.1

where $C(L,L_\infty)$ stands for the cumulated catch of fish of length L and above. The slope estimated by the linear regression 4.4.6.1 is Z/K so that an estimate of Z is obtained from:

$$Z = K * \text{slope}$$

Table 4.4.6.1 and Fig. 4.4.6.1 show an application of Eq. 4.4.6.1 to *Upeneus vittatus* (the data used in the foregoing section). Note that the basic data also for this method is the sample - not the total catch (cf. Sections 4.4.4-4.4.5). The results of the Jones and van Zalinge analysis using the length groups 13-19 cm are similar to the results from the linearized catch curve analysis (Table 4.4.5.1). Apparently the Jones and van Zalinge method gives a narrower confidence interval than the linearized catch curve. The conclusion is now that Z is somewhere between 4.05-0.74 = 3.4 and 4.05+0.74 = 4.7.

Table 4.4.6.1 **Cumulated catch curve based on length composition data for the yellow striped goatfish, *Upeneus vittatus* from Manila Bay, Philippines (from Ziegler, 1979). L_∞ = 23.1 cm, K = 0.59 per year (compare Fig. 4.4.6.1)**

L1-L2	catch C(L1,L2)	cumulated catch $C(L1,L_\infty)$	$\ln(C(L1,L_\infty))$ Y	$\ln(L_\infty-L1)$ X	slope Z/K	remarks
6- 7	3	3816	8.25	2.84	-	not
7- 8	143	3813	8.25	2.78	-	used.
8- 9	271	3670	8.21	2.71	-	Not
9-10	318	3399	8.13	2.65	-	under
10-11	416	3081	8.03	2.57	-	full
11-12	488	2665	7.89	2.49	-	exploi-
12-13	614	2177	7.69	2.41	-	tation
13-14	613	1563	7.35	2.31	-	portion
14-15	493	950	6.86	2.21	-	used
15-16	278	457	6.12	2.09	5.61	in the
16-17	93	179	5.19	1.96	6.20	regres-
17-18	73	86	4.45	1.81	5.98	sion
18-19	7	13	2.56	1.63	6.86	analy-
19-20	2	6	1.79	1.41	6.51	sis
20-21	2	4	1.39	1.13	-	not used
21-22	0	2	0.69	0.74	-	too close
22-23	1	2	0.69	0.10	-	to L_∞
23-24	1	1	0.00	-2.30	-	

Details of the five regression analyses:

L1-L2	slope*K Z	number of obs. n	Student's distrib. 95% t(n-2)	variance of slope sb^2	confidence limits Z ± Kt(n-2) sb
13-14	-	1	-	-	-
14-15	-	2	-	-	-
15-16	3.30	3	12.71	0.131	3.30 ± 2.71
16-17	3.66	4	4.30	0.118	3.66 ± 0.87
17-18	3.53	5	3.18	0.047	3.53 ± 0.41
18-19	4.05	6	2.78	0.202	4.05 ± 0.74
19-20	3.84	7	2.57	0.110	3.84 ± 0.50

Which method to use, the "linearized catch curve" or the "Jones and van Zalinge method" is at the moment a matter of taste. Perhaps future research will show that one is superior to the other. The main advantage of the linearized catch curve lies in its conceptual properties. The concept of the cumulated catch curve might be somewhat more complicated to grasp. Further, as will be demonstrated in Section 6.2 the set up of the linearized catch curve analysis is suitable also for the estimation of trawl mesh selection.

(See **Exercise(s)** in Chapter 17).

4.4.7 Summary of the linearized catch curve methods

Fig. 4.4.7.1 shows a summary of the versions of the linearized catch curve (Eq. 4.4.2.1) discussed in Section 4.4. These all originated from the catch equation (or Baranov's equation), Eq. 4.2.7. The following five models were introduced, three based on age composition data:

1) The linearized catch curve with constant time intervals. Eq. 4.4.3.1

2) The linearized catch curve with variable time intervals. Eq. 4.4.4.2 (generalizes Eq. 4.4.3.1)

3) The cumulated catch curve. Eq. 4.4.4.1

and two based on length composition data:

4) The linearized length converted catch curve. Eq. 4.4.5.3, derived from Eq. 4.4.4.2

5) The Jones and van Zalinge method. Eq. 4.4.6.1, derived from Eq. 4.4.4.1

4.5 Beverton and Holt's Z-equations

The method dealt with in this section first estimates Z from the mean length of the fish in the catch and the von Bertalanffy parameters K and L_∞. It is then shown that even L_∞ can be estimated together with an estimate of Z/K. Thus, the data requirements are even less than those for the length based linearized catch curve methods, for which the relative (not necessarily the absolute) size composition of the catch should be known (cf. Sections 4.4.4-6).

The derivation of Beverton and Holt's Z equations, however, is somewhat more complicated from a mathematical point of view. In the following we attempt to explain and justify the methods by aid of a numerical example, skipping the mathematics.

Assume that samples of length compositions of catches have been collected from a fishery in 1960, 1970 and 1980, and that the results are those shown in Table 4.5.0.1. The example used here is a hypothetical data set, that is, a data set for which the true parameters are known. The length groups represent only that part of the stock which is under full exploitation (cf. discussion in Section 4.4.5), the part for which Z is assumed to remain constant for all length groups. Z may of course be different for the three years in question.

In addition, the von Bertalanffy parameters K and L_∞ must be known. For the example in Table 4.5.0.1 the parameters are:

$K = 0.3$ per year , $L_\infty = 100$ cm

GENERAL LINEARIZED CATCH CURVE MODEL

$$\ln C(t1,t2) = d - Zt1 + \ln[1 - \exp(-Z(t2-t1))]$$

(Eq. 4.4.2.1)

<table>
<tr><td rowspan="2">Models based on age composition data</td><td colspan="2">$\Delta t = t2-t1$ is constant
LINEARIZED CATCH CURVE EQUATION WITH CONSTANT TIME INTERVALS
$\ln C(t,t+\Delta t) = g - Zt$
(Eq. 4.4.3.1)</td></tr>
<tr><td colspan="2">$\Delta t = t2-t1$ is variable</td></tr>
<tr><td></td><td>Δt = small
LINEARIZED CATCH CURVE EQUATION WITH VARIABLE TIME INTERVALS
$\ln \frac{C(t,t+\Delta t)}{\Delta t} = c - Z(t+\Delta t/2)$
(Eq. 4.4.4.2)</td><td>$\Delta t = \infty$
CUMULATED CATCH CURVE EQUATION
$\ln C(t,\infty) = d - Zt$
(Eq. 4.4.4.1)</td></tr>
<tr><td>Conversion of age into length</td><td>$\Delta t(L1,L2) = \frac{1}{K}\ln\left(\frac{L\infty-L1}{L\infty-L2}\right)$
(Eq. 4.4.5.1)
and
$t(L1) - \frac{t(L1)+t(L2)}{2} =$
$t_o - \frac{1}{K}\ln\left(1 - \frac{L1+L2}{2L\infty}\right)$
(Eq. 4.4.5.2)</td><td>$t(L) = t_o - \frac{1}{K}\ln\left(1-\frac{L}{L\infty}\right)$
(Eq. 3.3.3.3)</td></tr>
<tr><td>Models based on length composition data</td><td>**LINEARIZED LENGTH CONVERTED CATCH CURVE**
$\ln \frac{C(L1,L2)}{\Delta t(L1,L2)} = c-Z\,\frac{t(L1)+t(L2)}{2}$
(Eq. 4.4.5.3)</td><td>**JONES AND VAN ZALINGE METHOD**
$\ln C(L,L_\infty) = a + \frac{Z}{K}\ln(L_\infty-L)$
(Eq. 4.4.6.1)</td></tr>
</table>

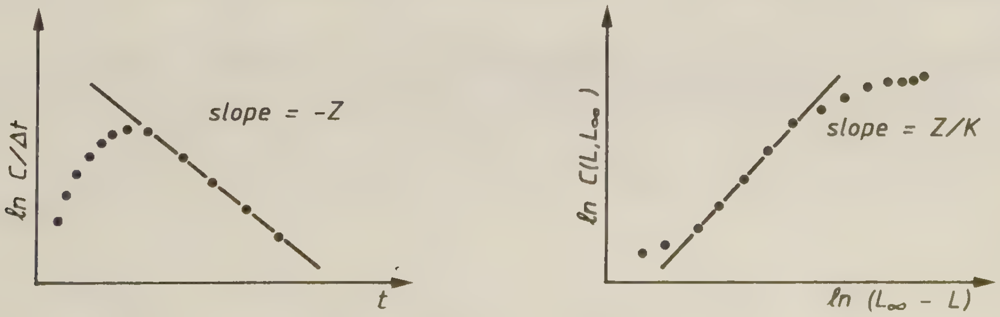


Fig. 4.4.7.1 Summary of models presented in Section 4.4

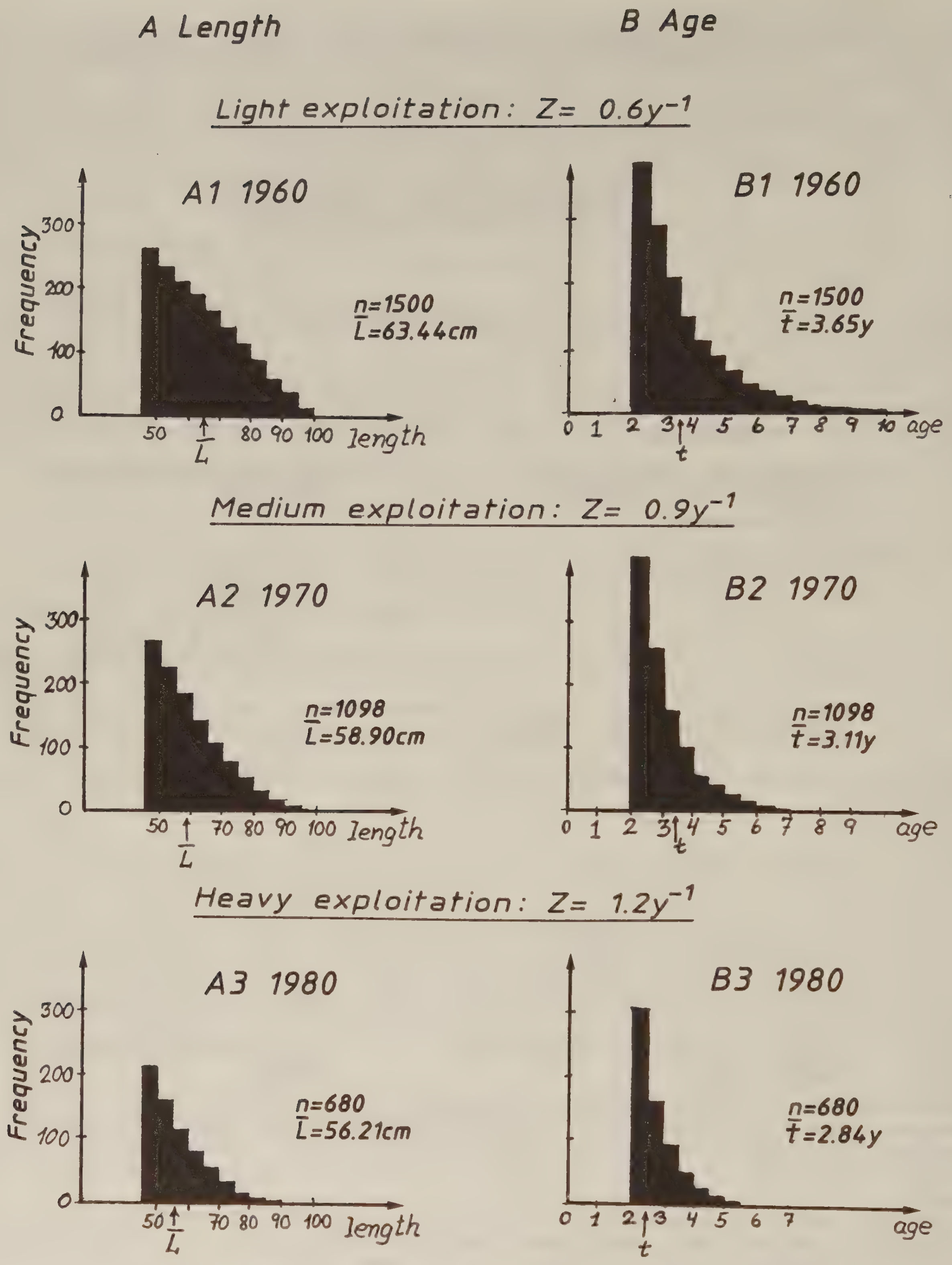


Fig. 4.5.0.1 Hypothetical example to illustrate Beverton and Holt's Z-equation (Eq. 4.5.0.1). K = 0.3, L' = 45 cm and L_∞ = 100 cm

Table 4.5.0.1 **Length composition samples of catches (hypothetical example to illustrate the method). Mean lengths are calculated as demonstrated in Section 2.1 (Table 2.1.2). L' is some length for which all fish of that length and longer are under full exploitation**

year			numbers caught 1960	 1970	 1980
	length group	mid-length			
L' =	45-50	47.5	256	268	212
	50-55	52.5	237	226	161
	55-60	57.5	211	180	116
	60-65	62.5	187	141	79
	65-70	67.5	161	105	52
	70-75	72.5	138	76	31
	75-80	77.5	113	50	17
	80-85	82.5	87	30	8
	85-90	87.5	62	15	3
	90-95	92.5	36	6	1
	95-100	97.5	12	1	0
Total (sample size)			1500	1098	680
Σ number * midlength			95155	64670	38220
mean length , $\bar{L}$			63.44	58.90	56.21

Figs. 4.5.0.1A1 to 4.5.0.1A3 show the length distributions. The mean length $\bar{L}$ has a decreasing trend in time. Thus in 1980 fewer fish survived to a certain length compared to 1970 and 1960. If this change in mean length is caused only by a change in the fishing mortality on the stock it follows that:

(F in 1960) < (F in 1970) < (F in 1980)

In other words, the lower the mean length in the catch, the higher is F. Under the assumption of a constant natural mortality, M, the same conclusion holds for the total mortality, Z = F+M, as for F.

4.5.1 Beverton and Holt's Z-equation based on length data

Beverton and Holt (1956) showed that the functional relationship between Z and $\bar{L}$ is:

$$Z = K \frac{L\infty - \bar{L}}{\bar{L} - L'} \qquad (4.5.1.1)$$

where $\bar{L}$ is the mean length of fish of length L' and longer, while L' is "some length for which all fish of that length and longer are under full exploitation". Note that L' is the lower limit of the corresponding length interval. For the choice of L', the same comments as those given in Section 4.4.2 and 4.4.6 can be applied. In the example the value of L' = 45 cm has been chosen.

The Z-values for the three years in the example become (compare Table 4.5.0.1 and Fig. 4.5.0.1A):

$$Z(1960) = 0.3\ \frac{100.0 - 63.44}{63.44 - 45.0} = 0.6 \text{ per year}$$

$$Z(1970) = 0.3\ \frac{100.0 - 58.90}{58.90 - 45.0} = 0.9 \text{ per year}$$

$$Z(1980) = 0.3\ \frac{100.0 - 56.21}{56.21 - 45.0} = 1.2 \text{ per year}$$

The method is refined in Section 4.5.4.

(See **Exercise(s)** in Chapter 17).

4.5.2 Beverton and Holt's Z-equation based on age data

As was the case for the linearized catch curve method the length based Beverton and Holt formula has an age based parallel. It is mentioned here mainly because it illustrates a basic feature about the Beverton and Holt theory (cf. Section 1.3).

The age compositions corresponding to the length compositions in Table 4.5.0.1 are shown in Table 4.5.2.1 and Fig. 4.5.0.1B. The same reasoning as was used for the mean length can be used for the mean age: the larger the fishing mortality the smaller is the mean age, $\bar{t}$, in the catch, or, the higher the mean age the smaller is F. This may be considered an implication of the functional relationship between age and length. However, the exact relationship between mean age and mean length is somewhat complicated. It should be noticed that the von Bertalanffy equation does not transform the mean age into the mean length, because the growth equation is not linear.

Table 4.5.2.1 Age composition of the samples given in Table 4.5.0.1

age group	1960	1979	1980
2.0- 2.5	390	399	308
2.5- 3.0	289	255	169
3.0- 3.5	214	162	92
3.5- 4.0	150	103	52
4.0- 4.5	118	66	28
4.5- 5.0	87	42	15
5.0- 5.5	64	27	8
5.5- 6.0	48	17	6
6.0- 6.5	35	11	2
6.5- 7.0	26	7	-
7.0- 7.5	19	4	-
7.5- 8.0	14	3	-
8.0- 8.5	11	2	-
8.5- 9.0	8	-	-
9.0- 9.5	6	-	-
9.5-10.0	4	-	-
10.0-10.5	3	-	-
10.5-11.0	2	-	-
11.0-11.5	2	-	-
sample size	1500	1098	680
Σnumber*age	5475	3415	1980
mean age, $\bar{t}$	3.65	3.11	2.84

The age based parallel to Eq. 4.5.1.1 is somewhat simpler. Beverton and Holt showed that:

$$Z = \frac{1}{\bar{t} - t'} \tag{4.5.2.1}$$

where $\bar{t}$ is the mean age of all fish of age t' and older, and where t' is "some age for which all fish of that age and older are under full exploitation". For the example we get that the age, t', corresponding to L' = 45 cm is 2 years ($2 = (1/K)\ \ln[(L_\infty - 45)/L_\infty]$), and

$$Z(1960) = \frac{1}{3.65 - 2.0} = 0.6 \text{ per year}$$

$$Z(1970) = \frac{1}{3.11 - 2.0} = 0.9 \text{ per year}$$

$$Z(1980) = \frac{1}{2.84 - 2.0} = 1.2 \text{ per year}$$

4.5.3 Beverton and Holt's Z-equation based on length at first capture

There exist a third version of the equation which mathematically is equivalent to Eq. 4.5.1.1, but conceptually different. Consider a "gear selection curve" for, say, a trawl as shown in Fig. 4.5.3.1. The vertical axis shows the percentage of the fish entering the gear (trawl) that is retained by the meshes. Lc, or L50%, is the length at which 50% are retained and 50% escape through the meshes. Note that Lc < L'. (In Chapter 6 we shall further discuss properties of selection curves).

This version of the Beverton and Holt Z-equation reads:

$$Z = K\,\frac{L\infty - \overline{Lc}}{\overline{Lc} - Lc} \tag{4.5.3.1}$$

where Lc is the "length at which 50% of the fish entering the gear are retained" and $\overline{Lc}$ is the "average length of the entire catch". Eq. 4.5.3.1 has proved useful in many cases where estimates of Lc and $\overline{Lc}$ are available, but $\bar{L}$ and L' are not known.

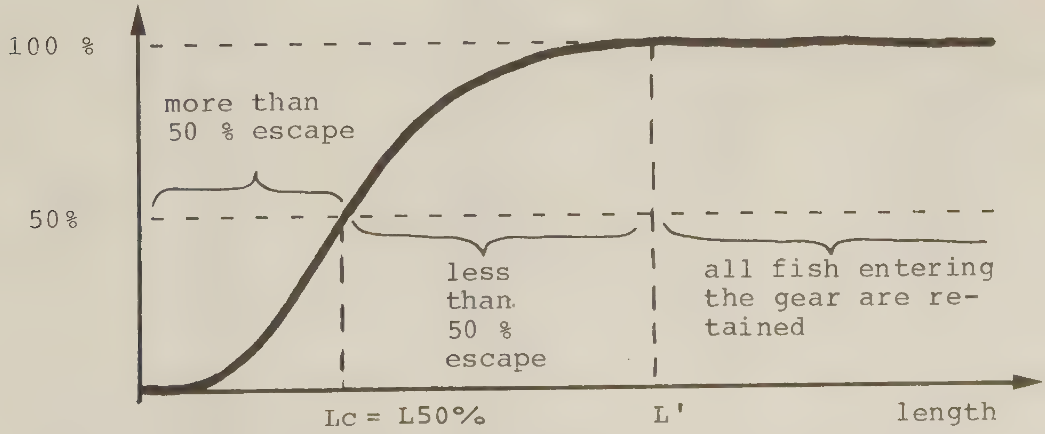


Fig. 4.5.3.1 Gear selection curve

4.5.4 The Wetherall et al. method

Wetherall, Polovina and Ralston (1987) suggested a special application of Eq. 4.5.1.1 by which L_∞ and Z/K can be estimated. As L' can take any value equal to and above the smallest length under full exploitation, Eq. 4.5.1.1 can give a series of estimates of Z, namely one for each choice of L'. This makes it possible to turn Eq. 4.5.1.1 into a regression analysis with L' as the independent variable. A series of algebraic manipulations shows that Eq. 4.5.1.1 is equivalent to:

$$\bar{L} - L' = a + b * L' \tag{4.5.4.1}$$

where

$$Z/K = -(1+b)/b \quad \text{and} \quad L_\infty = -a/b \tag{4.5.4.2}$$

or $b = -K/(Z+K)$ and $a = -b * L_\infty$

Thus, plotting $\bar{L} - L'$ against L' gives a linear regression from which a and b can be estimated and hence L_∞ and Z/K (see Fig. 4.5.4.1). Actually, Whetherall, Polovina and Ralston (1987) suggested a slightly different expression, so we call Eq. 4.5.4.1 the "modified Whetherall et al. me-

Table 4.5.4.1 Modified Wetherall et al. method (compare Fig. 4.5.4.2). The same data are used in Table 6.2.1

length interval L1-L2 X = L1	number caught C	mean length of fish longer than L', $\bar{L}$	$\bar{L} - L'$ Y	
3- 5	37	21.101	18.101	not used in analysis
5- 7	56	21.291	16.291	
7- 9	86	21.552	14.552	
9-11	129	21.917	12.917	
11-13	188	22.419	11.419	
13-15	258	23.101	10.101	
15-17	319	23.998	8.998	
17-19	352	25.108	8.108	
19-21	351	26.394	7.394	
21-23	324	27.801	6.801	
23-25	283	29.279	6.279	used in analysis
25-27	239	30.792	5.792	
27-29	196	32.320	5.320	
29-31	158	33.852	4.852	
31-33	123	35.392	4.392	
33-35	93	36.926	3.926	
35-37	69	38.447	3.447	
37-39	48	39.982	2.982	
39-41	31	41.516	2.516	
41-43	18	43.032	2.032	
43-45	10	44.462	1.462	not used
45-47	3	46.000	1.000	

Intercept a: 11.671 Confidence limits: [11.64 , 11.70]
Slope b : -0.2349 Confidence limits: [-.2359 , -.2340]
$L_\infty = -a/b = 49.7$ $K/Z = -(1+b)/b = 3.26$

thod". This method is especially suitable for situations where little or nothing is known about the fish stock in question. The estimation of L_∞ is especially useful.

It should be remembered that like the Beverton and Holt formula (Eq. 4.5.1.1) the method is based on the assumption of a constant parameter system which reduces its applicability.

An example of the modified Wetherall *et al*. method is given in Table 4.5.4.1 and Fig. 4.5.4.2. The calculation of $\bar{L}$ starts with the largest length class, in this case with L' = 45 cm.

The program "BHZWET" in the LFSA package of microcomputer programs (Sparre, 1987) can execute the modified Wetherall *et al*. analysis as well as the estimation of Z from Beverton and Holt's Z-equation. The "COMPLEAT ELEFAN" package of microcomputer programs (Gayanilo, Soriano and Pauly, 1988) also contains such programs.

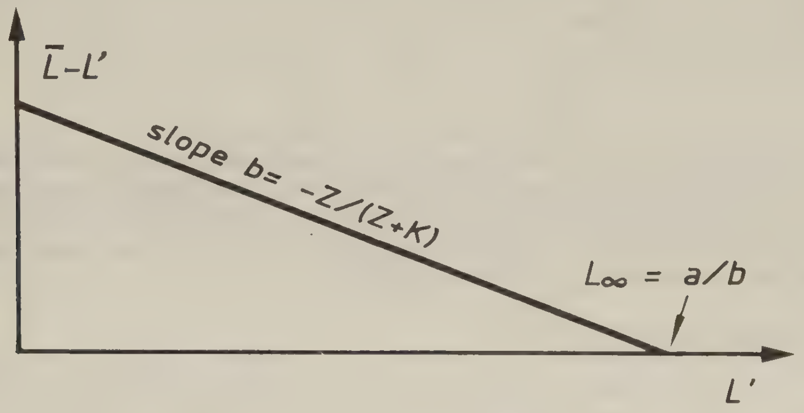


Fig. 4.5.4.1 Modified Wetherall *et al*. plot

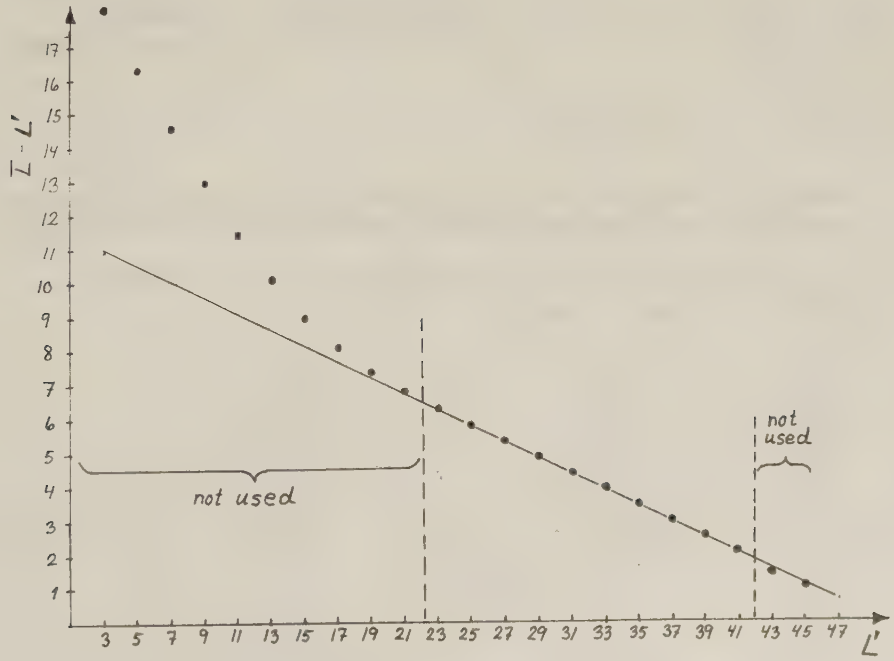


Fig. 4.5.4.2 Modified Wetherall *et al*. plot based on data from Table 4.5.4.1

4.6 A plot of Z on effort for separate estimates of F and M

The estimate of Z derived by the methods described in the foregoing section can sometimes be used to obtain estimates of F, the fishing mortality and M, the natural mortality. This is possible if pairs of estimates of Z and effort are available for a number of time periods covering a wide range of efforts.

Let Z(y) and f(y) be the total mortality and the effort in year y. The method (Paloheimo, 1958, 1961 and 1980) is based on Eq. 4.2.3: Z = F + M, and the assumption that effort and fishing mortality are related in the simple manner of Eq. 4.3.0.7: F(y) = q * f(y), where q is the catchability coefficient. Inserting Eq. 4.3.0.7 into Eq. 4.2.3 gives:

$$Z(y) = M + q\ f(y) \qquad (4.6.1)$$

With Z as the dependent variable and f as the independent variable Eq. 4.6.1 becomes a linear regression with slope q and the natural mortality, M, as the intercept. The fishing mortality for the time period y, F(y), is derived from:

$$F(y) = Z(y) - M$$

Table 4.6.1 and Fig. 4.6.1 show an application of Eq. 4.6.1 to Selaroides leptolepis in the Gulf of Thailand (from Boonyubol and Hongskul, 1978). In this case Z is calculated by Eq. 4.5.3.1, i.e. the observations are Lc, the 50% retention length, and $\overline{Lc}$, the mean length of all fish in the catch. Effort is given in units of millions of trawling hours per year. Fig. 4.6.1 shows the plot of Z on effort.

Using linear regression we find:

Slope : q = 0.244
Intercept: M = 2.05 per year

Confidence limits are calculated as described in Section 2.4:

$$sq^2 = \frac{1}{n-2}\left[\left(\frac{sy}{sx}\right)^2 - q^2\right] = \frac{1}{6}\left[\left(\frac{0.7518738}{2.7423475}\right)^2 - 0.2440881^2\right] =$$

$$= 0.002599$$

Table 4.6.1 Data for estimation of M and q for Selaroides leptolepis in the Gulf of Thailand (From Boonyubol and Hongskul, 1978)

year	effort [a)]	mean length	$Z = K\,\frac{L_\infty - \overline{Lc}}{\overline{Lc} - Lc}$	
y	f(y)	$\overline{Lc}$ cm		
1966	2.08	13.25	2.41	
1967	2.08	13.01	2.69	
1968	3.50	12.99	2.72	K = 1.16 per year
1969	3.60	13.07	2.62	Lc = 10.0 cm
1970	3.80	12.37	3.73	L_∞ = 20.0 cm
1971	-	- no data	-	
1972	7.19	12.30	3.68	
1973	9.94	12.01	4.61	
1974	6.06	12.60	3.30	

a) effort in units of millions of trawling hours

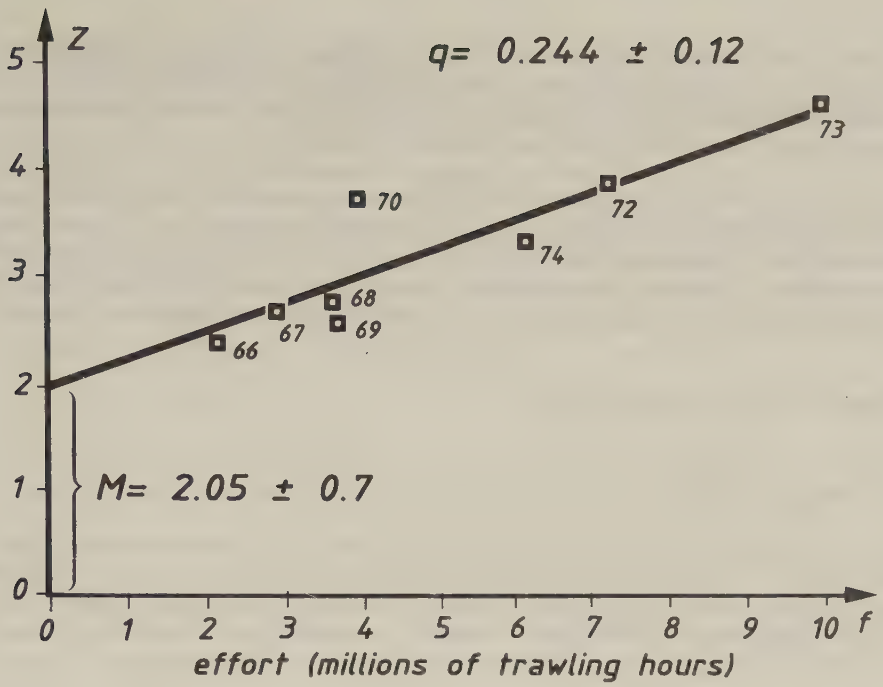


Fig. 4.6.1 Plot of total mortality, Z, on effort, f, for estimation of natural mortality, M, and catchability coefficient, q. Based on data from Table 4.6.1. Selaroides leptolepis in the Gulf of Thailand (from Boonyubol and Hongskul, 1978)

$$sM^2 = sq^2 \left(\frac{n-1}{n} sx^2 + \bar{x}^2 \right) =$$

$$0.002599 * \left(\frac{7}{8} 2.7423475^2 + 4.78125^2 \right) = 0.07650$$

The 95% confidence limits for q and M are

$$q: [q - sq * t(n-2) , q + sq * t(n-2)]$$

$$= [0.244 - 0.0510 * 2.45 , 0.244 + 0.0510 * 2.45]$$

$$= [0.12 , 0.68]$$

$$M: [M - sM * t(n-2) , M + sM * t(n-2)]$$

$$= [2.05 - 0.27662 * 2.45 , 2.05 + 0.27662 * 2.45]$$

$$= [1.3 , 2.7]$$

The estimates of M and q obtained this way are not precise. For example, they are based on the assumption that effort is proportional to fishing mortality, i.e, that q is constant - an assumption which always can be questioned due to an increasing trend in fishing efficiency. One trawling hour in 1970 may not create the same fishing mortality as a trawling hour in 1980 because fishing vessels in 1980 were often better equipped (more efficient gears, larger engines etc) than the vessels in 1970. If this is the case in the example of Fig. 4.6.1 the more recent observations should be moved to the right hand side to compensate for the increased efficiency. This would correspond to expressing the effort in units of 1966 trawling hours per year. The effect on the regression analysis would be a smaller slope and a larger intercept, i.e. the estimate of M would be larger and that of q smaller.

Note also that the application of the method is only possible if there is a clear change in effort during the period considered (cf. Exercise 4.6). If effort (the independent variable) remains constant during the period we should not be able to fit a line to the points.

In Section 4.7 we shall continue the discussion of natural mortality, and two alternative simple methods for estimation of M will be suggested. All simple methods for estimation of natural mortality are questionable and it is recommended always to assess the reasonability of any estimate by comparing results of alternative methods and estimates of M for similar species.

(See **Exercise(s)** in Chapter 17).

4.7 Natural mortality

One method of estimating the natural mortality coefficent, M, was introduced in Section 4.6. Many other procedures have been suggested in the literature. Most of them, perhaps all, rank no higher than "guesstimates" or "qualified" guesses. Some of these methods are widely used and others may be so in the future. They therefore merit discussion.

Recall that the natural mortality is the mortality created by all other causes than fishing, e.g., predation including cannibalism, diseases, spawning stress, starvation, and old age. Predation and starvation mortalities and several others are linked to the ambient ecosystem. The same species may have different natural mortality rates in different areas dependent on the density of predators and competitors whose abundance is influenced by fishing activities.

As direct measurements of M are often imposible to obtain, it has been attempted to identify quantities which can be assumed proportional to M and which are easier to measure (or estimate).

The von Bertalanffy curvature parameter, K, has been demonstrated to be linked to the longevity of the fish (Beverton and Holt, 1959) and longevity (Tanaka, 1960; Holt, 1965 and Saville, 1977) is related to mortality (compare with Eq. 4.5.2.1). Other methods involving longevity are treated in Section 4.7.1.

As a rough generalization, fish species with a high K-value have a high M-value, and species with a low K-value have a low natural mortality. A slow growing species (low K) simply cannot bear high natural mortality - if it did it would soon go extinct. Beverton and Holt (1959) found that values of the ratio M/K mostly lie in the range of 1.5 to 2.5.

Natural mortality must also be linked to L_∞ or the maximum weight of the species, W_∞, since large fish have fewer predators than small fish. It has been suggested that M can be predicted from the body size for certain groups of animals (Taylor, 1960, for bivalves and Peterson and Wroblewski, 1984, for pelagic fish).

Rikhter and Efanov (Section 4.7.3) demonstrated that fish with a high natural mortality mature early in life, compensating for the high M by starting to reproduce earlier. Also can be mentioned that Gunderson and Dygert (1988) found a relationship of M with the ratio of gonad weight to somatic weight. This is reasonable because fish with a high mortality may compensate by producing more eggs.

Further, as most biological processes go faster at higher temperatures (of course within certain limits) one could imagine the natural mortality to be related to the environmental temperature. Pauly's empirical formula (Section 4.7.2) describes M as a function of K, L_∞ and T, the ambient temperature.

4.7.1 Natural mortality and longevity

Intuitively, we would consider longevity as something more closely related to mortality than K, L_∞ or ambient temperature. The basic concept of mortality coefficients was discussed in Section 4.2 (e.g. Fig. 4.2.1). Alagaraja (1984) suggested another way of illustrating the concept of the mortality coefficient. He tentatively defined the natural lifespan of fish species (or the longevity) as the age at which 99% of a cohort had died if it had been exposed to natural mortality only (i.e. if Z = M). If Tm stands for longevity and $M_{1\%}$ stands for the natural mortality corresponding to a 1 percent survival, then:

$$M_{1\%} = - \ln(0.01)/Tm \tag{4.7.1.1}$$

The table below shows a collection of M-values defined by the 99% lifespan as well as the corresponding 99.9% lifespan.

Tm	(months)	3	4	5	6	7	8	10	12	
$M_{1\%}$	per year	18	14	11	9	8	6.9	5.5	4.6	
$M_{0.1\%}$	per year	28	21	17	14	12	10	8.3	6.9	
Tm	(years)	2	3	4	5	6	8	10	15	20
$M_{1\%}$	per year	2.3	1.5	1.2	.92	.77	.58	.46	.31	.23
$M_{0.1\%}$	per year	3.5	2.3	1.7	1.4	1.2	.86	.77	.69	.63

The above table can be used for a first rough evaluation of a mortality estimate in conjunction with a modal progression analysis (in particular for tropical species). If, for example, from the modal progression analysis (cf. Section 3.4.2) you find significant numbers (more than 1%) of three years old fish in the catch samples you should not accept a total mortality Z larger than 1.5. If you find very few (less than 1 per mille) you may accept a total mortality up to 2.3.

As the (natural) longevity usually is as difficult to observe as the natural mortality the relationship of mortalities and life span does not provide any easier way to estimate M, but it presents the concepts in a way which may be easier to grasp. How easy it is to estimate the longevity of a certain species depends on how diffucult it is to determine the age of the oldest specimens. If the age of the oldest specimens can be read from the hard parts and if the stock is unexploited it is relatively easy to get an estimate of Tm.

The above aspects were discussed by Hoenig (1983), who also developed a model for estimation of Z from observations on longevity. This model is basically the same as the Alagaraja (1984) model, but applies to Z rather than to M. Hoenig gave the following equation based on observations of Z and Tm and linear regression analysis:

$$\ln(Z) = a + b \ln(Tm) \quad \text{with}$$

$a = 1.46$ and $b = -1.01$ for fish
(based on 84 stocks of 53 species) and

$a = 1.23$ and $b = -0.832$ for molluscs
(based on 28 stocks of 13 species)

The above results were based mainly on data from unexploited stocks. Hoenig suggests Tm to be estimated from the mean age of "the oldest specimens" in a sample. This approach, however, is somewhat problematic as the estimate of Tm then depends on the sample size and on the proportion of the sample that is considered to represent "the oldest specimens".

Finally it should be emphasized that the often made assumption, that natural mortality remains constant for all age (size) groups within a species, is highly unrealistic for certain age (size) groups. Naturally, a small specimen is exposed to a larger predation mortality than a large one, simply because small animals have more predators than large animals. This difference in predation mortality (which is perhaps the dominating source of natural mortality for small fish) may be quite big.

When estimating predation mortalities from multispecies cohort analysis combined with work on stomach contents, the ICES Multispecies Working Group (ICES, 1986; Gislason and Sparre, 1987), found a factor of about 10 between the mortalities of 0-group fish and 2-group fish in the North Sea. (Multispecies VPA is further discussed in Section 10.2).

4.7.2 Pauly's empirical formula

Pauly (1980c) made a regression analysis of M (per year) on K (per year), L_∞ (cm) and T (average annual temperature at the surface in degrees Centigrade), based on data from 175 different fish stocks, and estimated the empirical linear relationship:

$$\ln M = -0.0152 - 0.279 \ln L_\infty + 0.6543 \ln K + 0.463 \ln T \qquad (4.7.2.1)$$

Table 4.7.2.1 shows values of M (per year) calculated by Pauly's formula for various combinations of L_∞, K and T. It should be kept in mind that Eq. 4.7.2.1 gives an estimate of M belonging to the category of "qualified guesses".

Eq. 4.7.2.1 may be correct for the "average fish", but may be way off the mark for any particular fish stock. The formula indicates that:

1. Small fish have high natural mortalities
2. Fast growing species have high natural mortalities
3. The warmer the ambient water the higher the natural mortality

The implication is that any other aspect which may influence the natural mortality is considered "random noise around the regression line" (Eq. 4.7.2.1). "Other aspects" are for example the behaviour (schooling, pelagic/demersal), the reproduction physiology, the ecosystem (abundance of predators). Pauly (1983) suggests to account for schooling in Eq. 4.7.2.1 by multiplication by 0.8 so that for schooling species the estimate becomes 20 percent lower:

$$M = 0.8 * \exp[-0.0152-0.279 * \ln L_\infty + 0.6543 * \ln K + 0.463 * \ln T\ ]$$

It should also be kept in mind that the quality of the input data to the regression analysis from which Eq. 4.7.2.1 was derived, can be questioned. to arrive at Eq. 4.7.2.1 Pauly needed "observations" of M, which as discussed in the foregoing are difficult to obtain. Any M value which was used for the estimation of Eq. 4.7.2.1 can be questioned. Therefore it is recommended only to use one decimal when presenting the result of Eq. 4.7.2.1. For some species the result seems to be twice or half of what it should be. However, when no other information on M is available as is most often the case one should not hesitate to apply Eq. 4.7.2.1 or the Rikhter and Efanov formula discussed in Section 4.7.3. Usually, it is only for unexploited stocks (M = Z) that we are able to estimate M.

Note that Pauly's formula is based on data for fish stocks only, and that the result depends on how you measure the length of the fish (total length, fork length, standard length etc.). As Pauly's equation gives only a rough estimate of M we can however ignore this detail. Do not use Pauly's formula for crustaceans, molluscs or cephalopods or any other non-fish animals as the formula does not cover these groups.

Table 4.7.2.1 **Natural mortality (per year) calculated by Pauly's formula for various combinations of L_∞, K and T**

	T = 5 C°				T = 25 C°			
L_∞ K	0.1	0.5	1.0	2.0	0.1	0.5	1.0	2.0
10	0.24	0.7	1.1	1.7	0.51	1.5	2.3	3.6
80	0.14	0.38	0.6	1.0	0.29	0.8	1.3	2.0
200	0.10	0.30	0.47	0.7	0.22	0.6	1.0	1.6

4.7.3 Rikhter and Efanov's formula

Beverton and Holt (1959) investigated a relationship between longevity, Tm, and the Lm/L_∞ ratio, where Lm is the length at first spawning. Holt (1962) noted that it was generally accepted that the Lm/L_∞ ratio was about 2/3 (average 0.64 with minimum and maximum values of 0.3 and 0.9).

Continuing these investigation in comparative dynamics for high latitude stocks, Rikhter and Efanov (1976) showed a close association between M and T_{m50}, the age when 50% of the population is mature (also called "the age of massive maturation"):

$$M = 1.521/(T_{m50}{}^{0.720}) - 0.155 \text{ per year} \qquad (4.7.3.1)$$

They also suggested that T_{m50} be equal to the "optimum age" defined as the age at which the biomass of a cohort is maximal.

Eq. 4.7.3.1 gives these values for T_{m50}:

T_{m50}	4 months	8 months	1 year	5 years	10 years
M	3.2	1.9	1.4	0.3	0.1

5 VIRTUAL POPULATION METHODS

This chapter deals in principle with the same problems as Chapter 4. However, the analysis and the data requirements are more detailed. The methods of Chapter 4 could be applied to data sets originating from small samples of the commercial catch or from research vessel catches. The methods of Chapter 5 require estimates of the total numbers caught by commercial fishing. Section 5.4 deals with the collection and processing of such data.

The methods are known as virtual population analysis or cohort analysis. The latter, because a cohort is followed from the moment it is recruited to the fishery until it goes extinct, which happens sooner the higher the fishing effort. The expression "virtual population" may be seen in contrast to a sample of a population. Information is required on how much there virtually was fished, in terms of numbers of fish. The total landings must be distributed over age groups (age based methods) or length groups (length based methods). The totals are raised from the age or length distributions of random samples of the landings, using information on the total landing in tonnes. Not until tables of the total landings in numbers, by age or length, per year or month, have been prepared, can we begin an analysis of the virtual population.

The idea behind these methods is that the total landings from a cohort in its lifetime is the first estimate of the numbers of recruits from that cohort. It is however, an underestimate because some fish must have died from natural causes. Given an estimate of M we can do a backwards calculation and find out how many fish belonging to the cohort were alive year by year and ultimately, how many recruits there were. At the same time we learn the values of the fishing mortality coefficient F, because we have calculated the numbers alive and know from the beginning how many of them were caught in any particular year.

The description above deals with the past history of cohorts and of the development of the fishing effort expressed by changes in F. The methods are also useful for describing the future effects of changes in the effort. The predictive versions of the analysis of virtual populations are also called Thompson and Bell methods (Sections 5.3 and 5.5).

5.1 Virtual population analysis (VPA)

This is the "historic" or past history cohort method. The calculations are complicated which gave rise to the introduction of a simplified version described in the next section as Pope's cohort analysis. VPA was developed by Fry (1949) and applied by Gulland (1965). It is reviewed by Jones (1984) and Pauly (1984), among others.

The data given in Table 4.4.3.1 are of the kind required for virtual population analysis, VPA, i.e. the total number caught by age group by the entire commercial fishery. To explain the concepts of VPA, consider the 1974-cohort of whiting caught in the North Sea as given in Table 4.4.3.1 (and Exercise 4.4.3).

The numbers caught (in units of millions) were:

$C(1974,0,1) = 599$, number caught between age 0 and age 1
$C(1975,1,2) = 860$, number caught between age 1 and age 2
$C(1976,2,3) = 1071$, number caught between age 2 and age 3
$C(1977,3,4) = 269$, number caught between age 3 and age 4
$C(1978,4,5) = 69$, number caught between age 4 and age 5
$C(1979,5,6) = 25$, number caught between age 5 and age 6
$C(1980,6,7) = 8$, number caught between age 6 and age 7

Suppose we know that the natural mortality was M = 0.2 per year for all age groups, and suppose further that we know that the fishing mortality for the 6-group in 1980 was F(1980,6,7) = 0.5 per year. With this knowledge it can be calculated how many survivors, N(1980,6), there must have been in the sea on 1 January 1980 to account for a catch of 8 million whiting in 1980. According to the catch equation (Eq. 4.2.7) N(1980,6) must satisfy:

$$C(1980,6,7) = N(1980,6)\ \frac{F}{Z}\ [1 - \exp(-Z(7-6))] \quad \text{or}$$

$$N(1980,6)\ \frac{0.5}{0.5+0.2}\ [1 - \exp(-(0.5+0.2) * 1.0)] = N(1980,6) * 0.36 = 8$$

$$\text{thus:}\quad N(1980,6) = \frac{8}{0.36} = 22.2 \text{ million}$$

Now that we know the number of survivors at the beginning of 1980 (which is equal to the number at the end of 1979), we can calculate how many whiting there must have been in the sea on 1 January 1979 to account for the catch C(1979,5,6) = 25 million.

We can also calculate which fishing mortality corresponds to this catch. To do so, Eq. 4.2.7 is used agin, but now in conjunction with the exponential decay model (Eq. 4.2.6):

$$C(1979,5,6) = N(1979,5)\ \frac{F}{Z}\ [1 - \exp(-Z)] \qquad (5.1.1)$$

$$N(1980,6) = N(1979,5) \exp(-Z) \qquad (5.1.2)$$

Eq. 5.1.2 is equivalent to:

$$N(1979,5) = N(1980,6) \exp(Z)$$

Inserting the value calculated above gives:

$$N(1979,5) = 22.2 * \exp(Z)$$

Inserting this and the number caught, C(1979,5,6) = 25 into Eq. 5.1.1 gives:

$$25 = 22,2 * \exp(Z)\ \frac{F}{Z}\ (1 - \exp(-Z))$$

which is equivalent to

$$\frac{25}{22.2} = \frac{F}{Z}\ (\exp(Z) - 1)$$

Inserting Z = F + M = F + 0.2 (M is assumed to be known) gives:

$$1.126 = \frac{F}{F+0.2}\ (\exp(F+0.2) - 1)$$

We have thus obtained an equation with F as the only unknown variable. Solving it gives an estimate of F. The equation above, however, is not the type of equation which can be solved by algebraic manipulations. It must be solved by some trial and error method. We shall later discuss how this minor technical problem can be circumvented, but for the time being we forget about it and note that F = 0.696 is the solution, i.e.:

$$1.126\ \frac{0.696}{0.696+0.2}\ (\exp(0.696+0.2) -1)$$

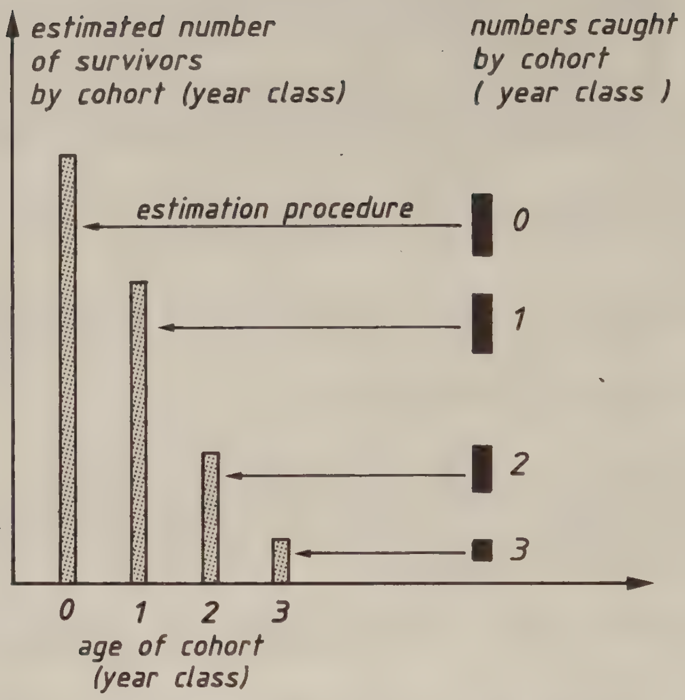

estimated number
of survivors
by cohort (year class)
numbers caught
by cohort
(year class)
estimation procedure
0
1
2
3
0
1
2
3
age of cohort
(year class)

Fig. 5.1.1 The basic features of VPA. For further explanation, see text

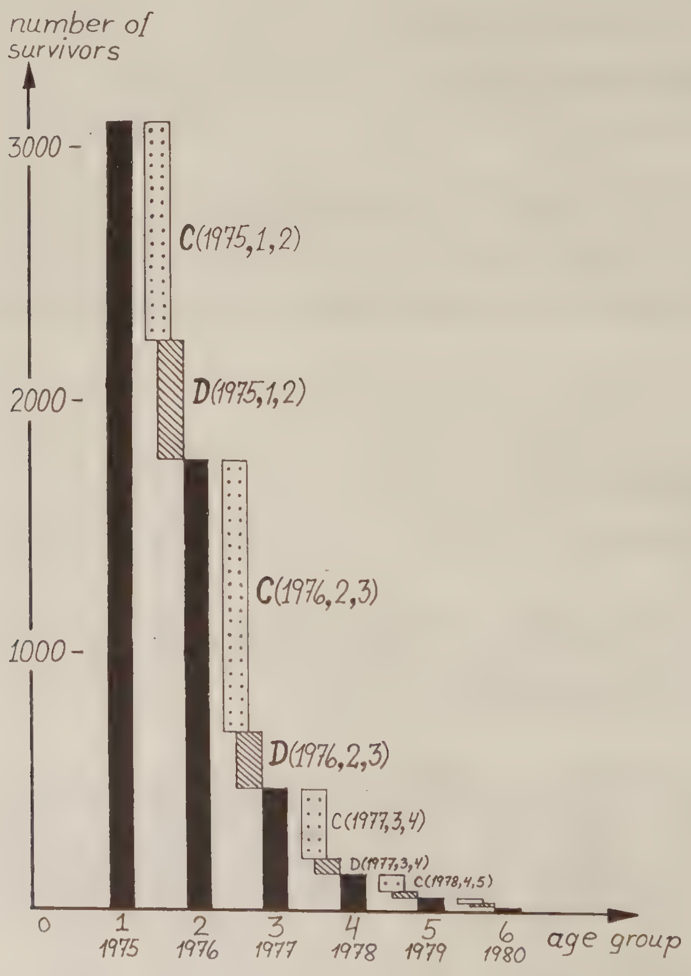


Fig. 5.1.2 Number of survivors, N (black bars), number caught $C = F \bar{N}$ and number of natural deaths, $D = M \bar{N}$, as derived by VPA in Table 5.1.1 for North Sea whiting

With the estimate of F(1979,5,6) = 0.696 the number of age group 5 fish on 1 January 1979 is easily found by means of the decay model:

$$Z(1979,5,6) = M + F(1979,5,6) = 0.2 + 0.696 = 0.896$$

$$\begin{aligned} N(1979,5) &= N(1980,6)\ \exp(Z(1979,5,6)) \\ &= 22.2\ \exp(0.896) = 54.4 \text{ million} \end{aligned}$$

The results of the calculations made so far may be summarized as follows:

age group t	year y	number caught during year y C(y,t,t+1)	fishing mort. during year y F(y,t,t+1)	no. surviving on 1 Jan. year y N(y,t)
0	1974	599	↑	↑
1	1975	860		
2	1976	1071	↑	↑
3	1977	269		
4	1978	69	↑	↑
5	1979	25	0.70	54.4
6	1980	8	0.50 *)	22.2

*) assumed to be known in advance.

The next pair N(1978,4) and F(1978,4,5) can be calculated exactly as those for the year 1979. In this way we can work "backwards in time" estimating numbers of survivors and fishing mortalities for each age group (as indicated by the arrows).

Note that contrary to the catch curve methods, we do not assume F (and Z) to remain constant. Each age group may have a different F-value. This method thus provides a more detailed analysis of the population than any of the other methods presented so far. The general pair of VPA-equations read (cf. the above given derivation):

$$\frac{C(y,t,t+1)}{N(y+1,t+1)} = \frac{F(y,t,t+1)}{M + F(y,t,t+1)} [\exp(F(y,t,t+1)+M) - 1] \qquad (5.1.3)$$

$$N(y,t) = N(y+1,t+1)\ \exp(F(y,t,t+1)+M) \qquad (5.1.4)$$

For the year 1978 in the example we get:

$$\frac{C(1978,4,5)}{N(1979,5)} = \frac{69}{54.4} = 1.268$$

and F(y,t,t+1) = F(1978,4,5) = 0.757 is the solution to Eq. 5.1.3 found by trial and error.

$$\begin{aligned} N(1978,4) &= N(1979,5)\ \exp(F(1979,4,5)+M = \\ &= 54.4\ \exp(0.757+0.2) = 141.9 \end{aligned}$$

Repeating this procedure for the years 1977-1974 the estimates of fishing mortalities and stock numbers shown in Table 5.1.1 are obtained.

The basic features of VPA can be summarized as follows:

From observations on the numbers caught in each age group VPA estimates how many fish there must have been in the sea to account for that catch, under the assumption that natural mortality is known (see Fig. 5.1.1). If the catch constitutes a small fraction of the stock (i.e. if F is small) the estimation of the stock size becomes more uncertain. Thus, the higher the fishing mortality the more dependable is the VPA.

Table 5.1.1 Results of VPA for the 1974 whiting cohort. (Data from Table 4.4.3.1, numbers in millions)

age group t	year y	number caught during year y C(y,t,t+1)	fishing mort. during year y F(y,t,t+1)	number of survivors on 1 Jan. year y N(y,t)
0	1974	599	0.16	4390.
1	1975	860	0.37	3054.
2	1976	1071	1.11	1729.
3	1977	269	0.99	465.
4	1978	69	0.76	142.
5	1979	25	0.70	54.4
6	1980	8	0.50 *)	22.2

*) assumed to be known in advance.

Natural mortality, M, is assumed to be known from investigations independent of the VPA, but is actually unknown in most cases. The reliability of VPA is also dependent on the size of M relative to F. Often the estimate of M is rather a "guesstimate" (qualified guess), but if M is small compared to F it may not matter so much that M is not well estimated. What a "guesstimate of M" means was discussed in Section 4.7.

Fig. 5.1.2 illustrates the cohort dynamics as described by VPA for North Sea whiting (Table 5.1.1). In this case M is relatively small compared to F, which can be seen by comparing the number caught (cf. Eq. 4.2.8 with t2-t1=1):

$$C(y,t,t+1) = F(y,t,t+1)\,\bar{N}(y,t,t+1) \qquad (5.1.5)$$

to the number of natural deaths:

$$D(y,t,t+1) = M(y,t,t+1)\,\bar{N}(y,t,t+1) \qquad (5.1.6)$$

Fig. 5.1.2 presents the results as the number of survivors, N, but we could as well have chosen the fishing mortality, F, as the basic result, because:

F and C determine N and
N and C determine F

so that there is a one to one correspondence between N and F when C and M are known.

To calculate confidence intervals for estimates of parameters the number of observations must be larger than the numbers of parameters. The number of parameters in VPA (the N's and the F's for the oldest age group) equals the number of observations (the C's) plus one. Therefore it is not possible to calculate confidence limits for the estimates of the N's (or the F's).

The data used to illustrate VPA were obtained from direct age reading (otoliths). However, input data could have been derived from time series of length frequency data which were resolved into cohort components by e.g the Bhattacharya method (Section 3.4). This aspect is further discussed at the end of Section 5.4.3.

The VPA is a method to analyse historical data for estimation of population parameters. The ultimate use of such parameters is to determine the optimum fishing strategy, i.e. the array of F at age, or the so-called "fishing pattern", which in the long term gives the highest yield from the stock in question. To assess alternative (future) fishing strategies we require a counterpart to the VPA, namely a model which can predict the stock and the catch for various assumptions on the future fishing pattern. The "Thompson and Bell model" (Section 5.3) is the predictive version of the VPA.

Mesnil (1988) presents a package of microcomputer programs, ANACO, ("ANAlysis of COhorts" or "L'ANAlyse des COhortes") which can perform the VPA calculations as described above. The ANACO package also offers a number of additional options, for example sensitivity analysis. Also the "COMPLEAT ELEFAN" package (Gayanilo, Soriano and Pauly, 1988) contains routines for VPA analysis.

5.2 Age based cohort analysis (Pope's cohort analysis)

As derived from the catch equation, the VPA implied the solution of Eq. 5.1.3 by some numerical techniques (some trial and error method). This is a minor technical problem when one has access to a computer. However, the problem can be circumvented in an easy way, so that VPA can also be carried out on a pocket calculator. The version of VPA suitable for pocket calculators is the "cohort analysis" developed by Pope (1972). (Reviewed in Jones, 1984 and Pauly, 1984).

Cohort analysis is conceptually identical to VPA, but the calculation technique is simpler. It is based on the approximation illustrated in Fig. 5.2.1 which shows the dynamics of a cohort during one year.

Actually, the catch is taken continuously during the year, but in cohort analysis the approximation is made that all fish are caught on one single day. This day is chosen to be 1 July, i.e. when one half of the year has elapsed. Consequently in the first half year the fish suffer only natural mortality so the number of survivors on 1 July becomes:

$$N(y,t+0.5) = N(y,t) \exp(-M/2)$$

Then, instantaneously, the catch is taken and the number of survivors becomes:

$$N(y,t) \exp(-M/2) - C(y,t,t+1)$$

(see Fig. 5.2.1). This number of survivors then suffers natural mortality in the second half year and the number of survivors at the end of the year is:

$$N(y+1,t+1) = [N(y,t) \exp(-M/2) - C(y,t,t+1)] \exp(-M/2)$$

For convinience of calculation the equation is rearranged as

$$N(y,t) = [N(y+1,t+1) \exp(M/2) + C(y,t,t+1)] \exp(M/2) \qquad (5.2.1)$$

Note that the F that caused computational problems in the VPA equation does not occur here. To apply Eq. 5.2.1 to the whiting example we start the same way as for the VPA by assuming the F for the oldest age group (the so-called terminal F) to be known, F(1980,6,7) = 0.5, while M = 0.2, and calculate N(1980,6):

$$N(1980,6) = \frac{C(1980,6,7)}{(F/Z)(1-\exp(-Z))} = 22.2$$

Then Eq. 5.2.1 is applied to calculate N(1979,5):

$$N(1979,5) = [N(1980,6)\exp(M/2) + C(1979,5,6)]\exp(M/2) = [22.2 * 1.1052 + 25] * 1.1052 = 54.7$$

and we continue in the same way:

$$N(1978,4) = [N(1979,5)\exp(M/2) + C(1978,4,5)]\exp(M/2) =$$

$$= [54.7 * 1.1052 + 69] * 1.1052 = 143.1$$

$$N(1977,3) = [143.1 * 1.1052 + 269] * 1.1052 = 472.0$$

..... etc. ....

Thus, as for the VPA we work backwards in time estimating a new stock number at each step. The calculation procedure is given in the first half of Table 5.2.1 (with a less extensive notation).

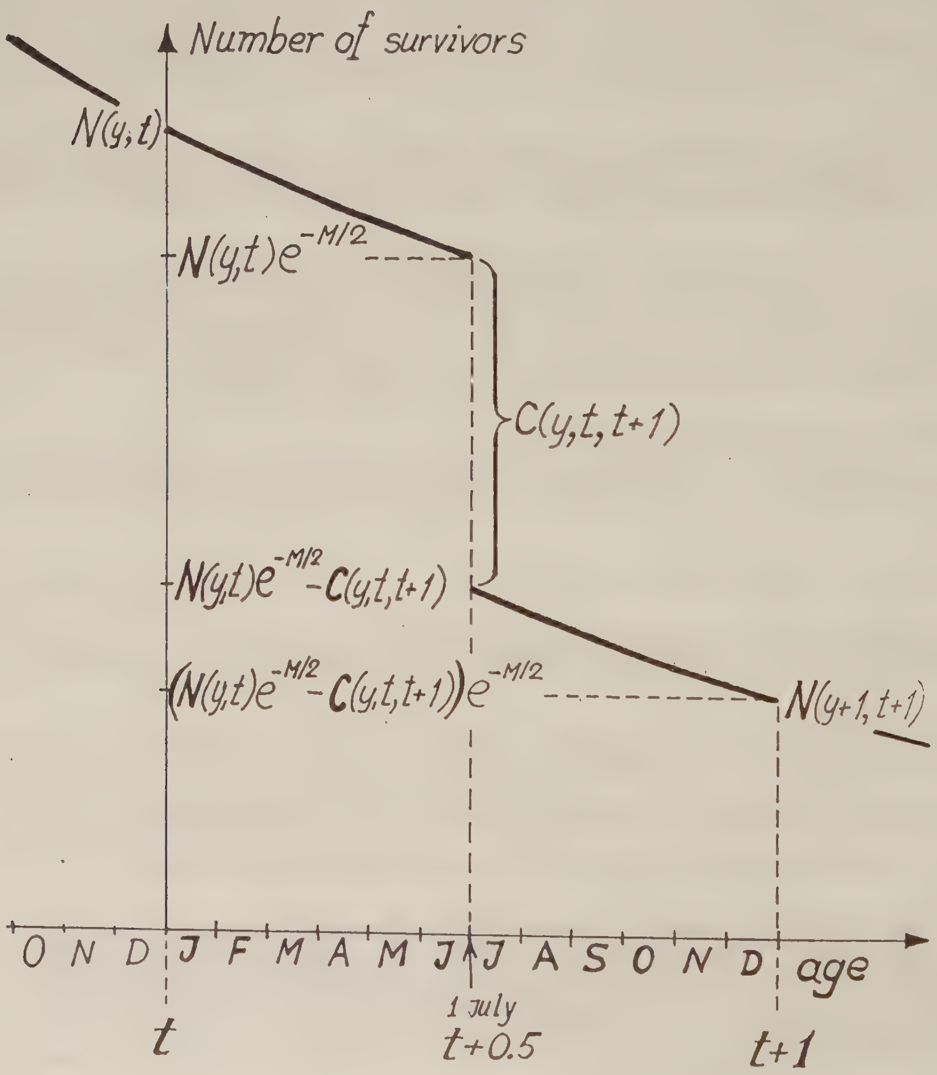

Number of survivors
N(y,t)
N(y,t)e^{-M/2}
C(y,t,t+1)
N(y,t)e^{-M/2}-C(y,t,t+1)
(N(y,t)e^{-M/2}-C(y,t,t+1))e^{-M/2}
N(y+1,t+1)
O N D J F M A M J J A S O N D
age
1 July
t
t+0.5
t+1

Fig. 5.2.1 Illustration of the approximation behind Pope's cohort analysis (for further explanation, see text)

Table 5.2.1 The computation procedure of Pope's age based cohort analysis, illustrated by the 1974-cohort of North Sea whiting (from Table 4.4.3.1)

POPE'S AGE BASED COHORT ANALYSIS

NATURAL MORTALITY FACTOR: $M = 0.2$ per year.

$$X = \exp(M/2) = \exp(0.2/2) = 1.1052$$

GUESS ON TERMINAL F: $F6 = 0.5$

STOCK NUMBERS:

$$N6 = \frac{8}{[0.5/(0.5+0.2)]\ [1 - \exp\{-(0.5+0.2)\}]}$$

$$N5 = (N6\ X + C5)\ X = (22.2\ X + 25)\ X = 54.7$$

$$N4 = (N5\ X + C4)\ X = (54.7\ X + 69)\ X = 143.1$$

$$N3 = (N4\ X + C3)\ X = (143.1\ X + 269)\ X = 472.1$$

$$N2 = (N3\ X + C2)\ X = (472.1\ X + 1071)\ X = 1760.3$$

$$N1 = (N2\ X + C1)\ X = (1760.3\ X + 860)\ X = 3100.4$$

$$N0 = (N1\ X + C0)\ X = (3100.4\ X + 599)\ X = 4448.9$$

FISHING MORTALITIES:

$$F6 = \text{(initial guess of terminal F)} = 0.50$$

$$F5 = \ln(N5/N6) - M = \ln(54.7/22.2) - 0.2 = 0.70$$

$$F4 = \ln(N4/N5) - M = \ln(143.1/54.7) - 0.2 = 0.76$$

$$F3 = \ln(N3/N4) - M = \ln(472.1/143.1) - 0.2 = 0.99$$

$$F2 = \ln(N2/N3) - M = \ln(1760.3/472.1) - 0.2 = 1.12$$

$$F1 = \ln(N1/N2) - M = \ln(3100.4/1760.3) - 0.2 = 0.37$$

$$F0 = \ln(N0/N1) - M = \ln(4448.9/3100.4) - 0.2 = 0.16$$

From the estimate of the N's the estimate of fishing mortality is obtained from:

$$F(y,t,t+1) = \ln \frac{N(y,t)}{N(y+1,t+1)} - M \qquad (5.2.2)$$

which follows from the exponential decay model solved for F:

$$N(y+1,t+1) = N(y,t) \exp(-F(y,t,t+1)-M) \qquad \text{(cf. Table 5.2.1)}$$

Comparing the VPA-results (Table 5.1.1) to the results of cohort analysis (Table 5.2.1) it appears that they differ, but the differences are small. Taking into consideration all the sources of uncertainty involved in this kind of calculation, one can say that the differences between VPA and cohort analysis are insignificant. Pope (1972) has shown that for $F < 1.2$

and $M < 0.3$ the differences will be small. The only advantage of the cohort analysis compared to VPA is that it can easily be carried out with a pocket calculator. To do the VPA on a pocket calculator requires the tedious trial and error way of calculating F (unless the calculator is fully programmable).

Eq. 5.2.1 was derived for a time period of one year. As for the catch curve (cf. Section 4.4.4) we may consider the catch during any time period, from t to $t + \Delta t$. In that case Eq. 5.2.1 should be replaced by the more general expression:

$$N(t) = [N(t+\Delta t)\exp(M\,\Delta t/2) + C(t,t+\Delta t)]\exp(M\,\Delta t/2) \qquad (5.2.3)$$

The derivation of Eq. 5.2.3 is similar to that of Eq. 5.2.1 and is shown in Fig. 5.2.2. The parallel to Eq. 5.2.2 is:

$$F(t,t+\Delta t) = \frac{1}{\Delta t}\ln\frac{N(t)}{N(t+\Delta t)} - M \qquad (5.2.4)$$

The year index, y, has disappeared in Eqs. 5.2.3 and 5.2.4, the main reason being that they usually are applied under the assumption of a constant parameter system (cf. Section 4.4.4). Further, when Δt varies it will not conform to the year intervals and the notation used for age groups is not suitable any more.

Eqs. 5.2.3 and 5.2.4 applied under the assumption of constant parameters is a typical application of cohort analysis for tropical fish stocks. They could be applied to all cohorts during one year or to the average annual catch during a sequence of years. We shall return to this in Section 5.4.

(See **Exercise(s)** in Chapter 17).

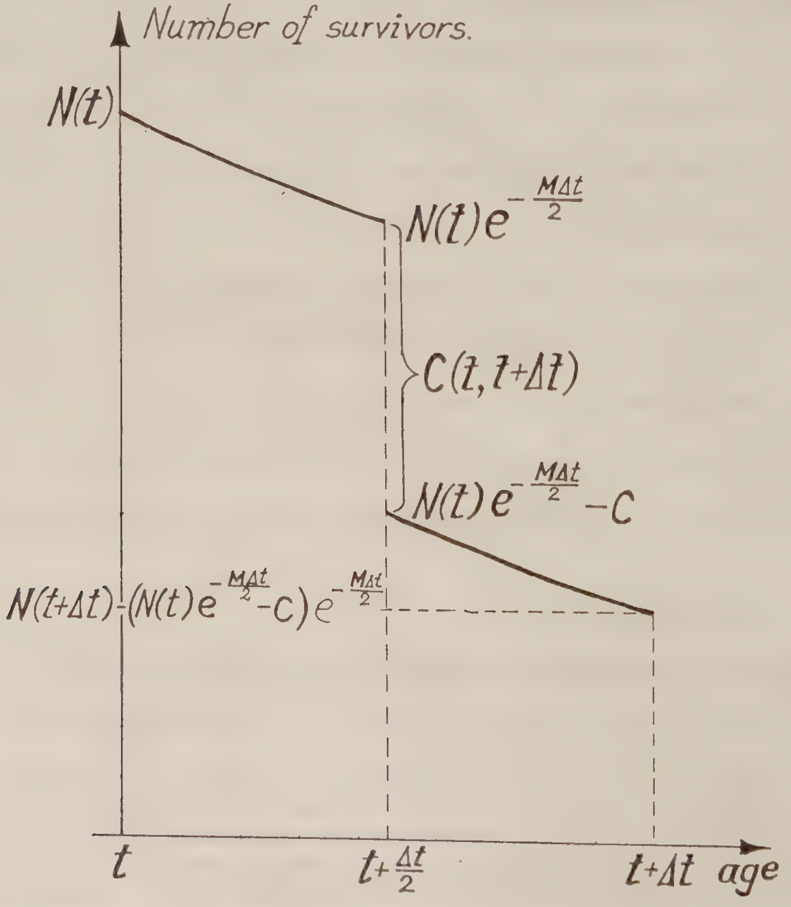

Number of survivors.
$N(t)$
$N(t)e^{-\frac{M\Delta t}{2}}$
$C(t,t+\Delta t)$
$N(t)e^{-\frac{M\Delta t}{2}}-C$
$N(t+\Delta t)=(N(t)e^{-\frac{M\Delta t}{2}}-C)e^{-\frac{M\Delta t}{2}}$
t
$t+\frac{\Delta t}{2}$
$t+\Delta t$
age

Fig. 5.2.2 Derivation of the formulas for Pope's cohort analysis in the case of variable time intervals (Eqs. 5.2.3 and 5.2.4)

5.3 A predictive age based model (Thompson and Bell)

VPA and Cohort Analysis (Section 5.1 and 5.2) are so-called historic models: they describe past events and estimate parameter values. They have a counterpart in the predictive models which, based on parameter estimates from the historic models, predict the development of a fishery, given assumptions on the future values of the effort (F) and mesh size (Tc) parameters. The predictive counterpart of VPA and Cohort Analysis was published by Thompson and Bell (1934). Widely used today, it was overlooked for many years because in 1934 only a few fisheries scientists had the mathematics required to appreciate it. In the forward version, here called the "*age based Thompson and Bell model*", we do not encounter the problem of VPA, that an iterative solution is required for one of the equations.

To illustrate the model we use an example on the Kuwait shrimp fishery (from Garcia and van Zalinge, 1982). Columns A to E in Table 5.3.1 contain the input. In this case the fishing mortalities, the F's, were estimated from biomass, estimated from the swept area method (cf. Chapter 13) and catch data. However, they might equally well have been estimated by Cohort Analysis or VPA.

The lifespan of the shrimp in question (*Penaeus semisulcatus*) is not much over one year, so the fishing mortalities given in Table 5.3.1 cover the major part of a cohort's life. Column C contains a "relative value" which is a quantity proportional to the price per kg of the unpeeled tails. As will be demonstrated later these prices are used to convert the yield per recruit in units of g/recruit into a unit of money/recruit. Note that the value increases as the shrimp grows larger such that grammes per recruit

Table 5.3.1 Age based Thompson and Bell model illustrated by data from the Kuwait shrimp fishery (from Garcia and van Zalinge, 1982). M = 3.0 per year for all ages

A	B	C	D	E	F	G	H	I	J	K
age t	wght. $w(t)$	value $v(t)$	fish. mort. $F(t)$	total mort. $Z(t)$	population $N(t)$	deaths $N(t)-N(t+\Delta t)$	catch $C(t)$	yield $Y(t)$	biomass $B(t)$	value $Y*v$
months	g	money unit	per year	per year	number	number	number	g	g	money unit
0	-	-	-	-	1000	-	-	-	-	-
1=Tr	5.7	0.73	1.20	4.20	705	295	84.2	480	4800	350.4
2	9.3	0.93	1.32	4.32	492	213	65.1	606	5500	562.7
3	13.0	1.20	1.32	4.32	343	149	45.5	591	5381	709.2
4	17.6	1.45	1.44	4.44	237	106	34.3	605	5042	877.3
5	22.0	1.70	1.92	4.92	157	80	31.2	687	4293	1167.9
6	26.1	1.90	1.20	4.20	111	46	13.1	343	3430	651.7
7	30.3	2.08	1.56	4.56	75.9	35.1	12.0	364	2797	757.1
8	33.8	2.14	1.20	4.20	53.5	22.4	6.4	216	2163	462.2
9	37.0	2.18	1.20	4.20	37.7	15.8	4.5	166	1665	361.9
10	40.3	2.23	1.80	4.80	25.3	12.4	4.7	189	1263	421.5
11	43.1	2.24	2.76	5.76	15.6	9.7	4.6	198	862	443.5
12	44.7	2.27	2.52	5.52	9.9	5.7	2.6	116	553	263.3
							Total:	4560	37749	7028.7
						Average biomass : 37749/12 = 3146				

are not proportional to the money return per recruit. Columns F - K illustrate the calculation procedure of the Thompson and Bell model. In column F we start with 1000 shrimps at age zero. Thus, in this case all results are relative to 1000 recruits. When cohort analysis (Section 5.2) has been carried out we have an estimate of the recruitment number, and we are in a position to estimate the actual yield and stock size. Here we do the calculation on a per 1000 shrimps basis, but we could have chosen any other unit (the results are relative anyway). We proceed as follows (compare the table):

Step 1: Calculate stock number at the beginning of each period:
$N(t+\Delta t) = N(t)\ \exp(-Z\Delta t)$
e.g. 1000 * exp(-4.20 * 0.0833) = 704.7 (age 1 month)
704.9 * exp(-4.32 * 0.0833) = 491.6 (age 2 months)

Step 2: Calculate total number of deaths in each period:
$N(t) - N(t+\Delta t)$
e.g. 1000 - 705 = 295 (age 1 month)
705 - 492 = 213 (age 2 months)

Step 3: Calculate numbers caught in each period:
$C(t) = (N(t)-N(t+\Delta t)) * F(t)/Z(t)$
e.g. 295 * 1.20/4.20 = 84.2 (age 1 month)
213 * 1.32/4.32 = 65.1 (age 2 months)

Step 4: Calculate weight of catch (yield) in each period:
$Y(t) = C(t) * w(t)$
e.g. 84.2 * 5.7 = 480 g (age 1 month)
65.1 * 9.3 = 605 g (age 2 months)

Step 5: Calculate stock biomass in each period:
$B(t) = Y(t)/(F(t) * \Delta t)$ (Δt = 1 month = 0.0833 year)
e.g. 480/(1.20 * 0.0833) = 4800 (age 1 month)
605/(1.32 * 0.0833) = 5500 (age 2 months)

Step 6: Calculate value of yield in each period:
$Y(t) * v(t)$
e.g. 480 * 0.73 = 350.4 (age 1 month)
605 * 0.93 = 562.7 (age 2 months)

Step 7: Calculate total yield, average biomass and total value
(see last row of Table 5.3.1)

As appears from the table the yield per 1000 shrimps is 4560 g with a value of 7029 units.

If the effort (or fishing mortality) is increased by, say, 20 percent the yield would become 4954 g and its value 7465. These figures are obtained by raising the fishing mortalities in column D of Table 5.3.1 by 20 percent, i.e. to use the F-array:

1.20 * 1.20 = 1.44, 1.32 * 1.20 = 1.58, ..., etc.

and then by going through the procedure described in the table once again. We say that we have applied the factor X = F/Fref = 1.20 to the reference array of fishing mortalities, i.e. the array given in column D. By applying a range of factors the yield and its value can be calculated for different effort levels (or different F-levels) relative to the values obtained for the reference level. A collection of such alternative yield calculations is shown in Table 5.3.2.

Fig. 5.3.1 shows the yield, value of yield and biomass curves corresponding to Table 5.3.2. Note that the value curve has a maximum whereas the yield curve has not in the range of factors (X = F/Fref) considered. When the price per kg varies with the individual size of the shrimps the two curves will have their maximum at different F-levels.

Table 5.3.2 Yield, value of yield and biomass for various F-levels. The reference F-array is given in column E of Table 5.3.1 (compare to Fig. 5.3.1)

X factor	total yield	total value	total biom.
0	0	0	5382
0.4	2549	4209	4271
0.8	4055	6396	3466
1.0	4560	7029	3146 *)
1.2	4954	7465	2870
1.5	5383	7842	2522
2.0	5814	8025	2075
3.0	6138	7683	1497

*) compare to Table 5.3.1

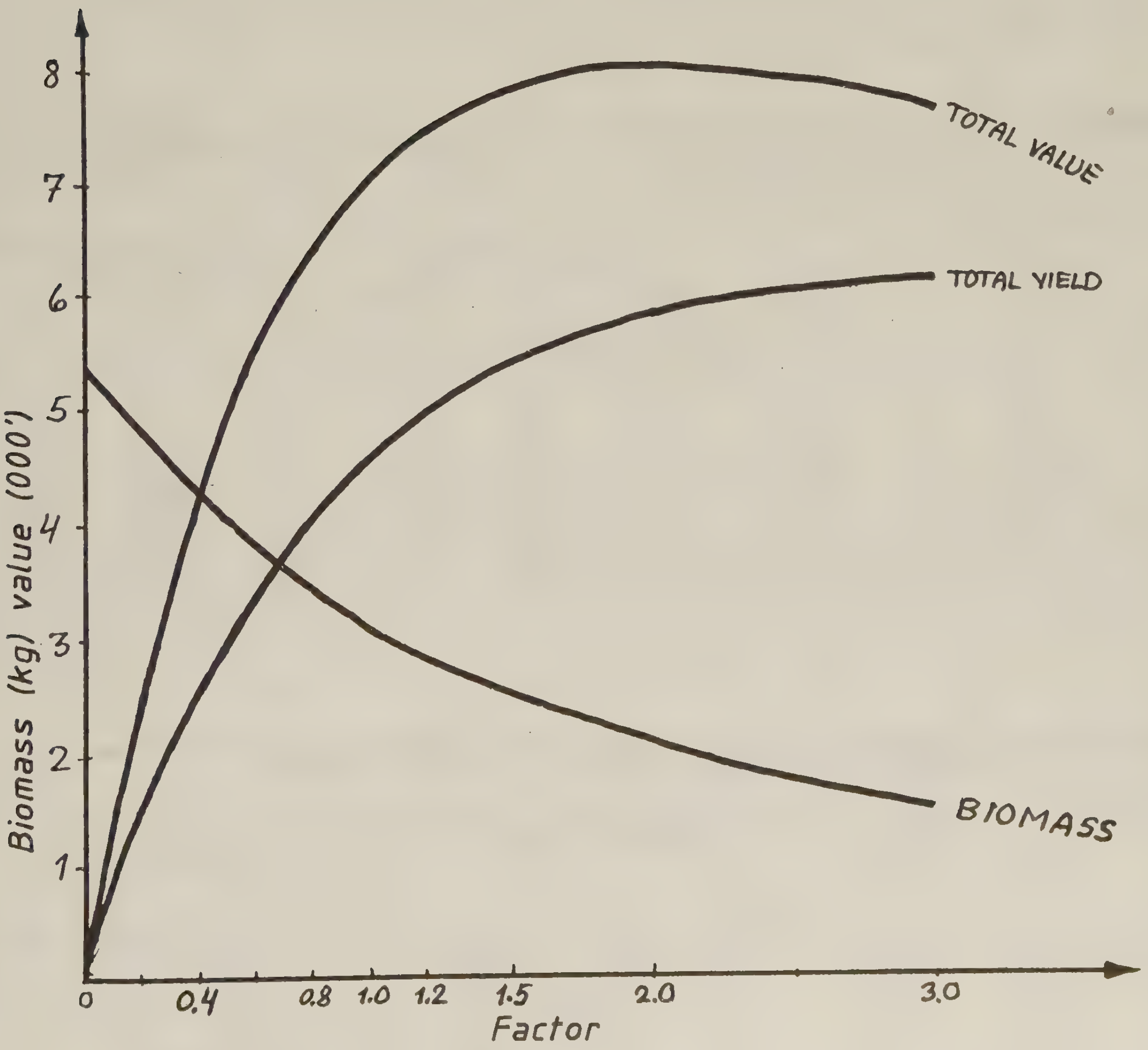


Fig. 5.3.1 Yield, value of yield and biomass per 1000 shrimps calculated by the age based Thompson and Bell model. Based on data in Table 5.3.2

Table 5.3.3 **Partitioning of total fishing mortality from the Kuwait shrimp fishery into an artisanal and an industrial component (from Garcia and van Zalinge, 1982)**

age (months)	artisanal F	industrial F	total F
1	0.720	0.480	1.20
2	0.960	0.360	1.32
3	0.840	0.480	1.32
4	0.480	0.960	1.44
5	0.600	1.320	1.92
6	0.480	0.720	1.20
7	1.080	0.480	1.56
8	0.480	0.720	1.20
9	0.084	1.116	1.20
10	0.120	1.680	1.80
11	0.240	2.520	2.76
12	0.240	2.280	2.52

Table 5.3.4 **Partitioning of yields and values of yield from Table 5.3.1 into an artisanal and an industrial component. The X-factors are as in Table 5.3.2**

total yield	total value	ARTISANAL FLEET			INDUSTRIAL FLEET		
		factor X1	yield	value	factor X2	yield	value
0	0	0	0	0	0	0	0
2549	4209	0.4	1048	1531	0.4	1501	2678
4055	6396	0.8	1773	2486	0.8	2284	3910
4560	7029	1.0	2048	2814	1.0	2512	4215
4954	7465	1.2	2281	3073	1.2	2673	4392
5383	7842	1.5	2563	3354	1.5	2819	4488
5814	8025	2.0	2903	3627	2.0	2910	4398
6138	7683	3.0	3291	3783	3.0	2847	3900

Table 5.3.5 **Assessment of the effect of varying the industrial effort while the artisanal effort is kept constant**

total yield	total value	ARTISANAL FLEET			INDUSTRIAL FLEET		
		factor X1	yield	value	factor X2	yield	value
2479	3603	1.0	2479	3603	0	0	0
3522	5403	1.0	2289	3250	0.4	1234	2154
4270	6598	1.0	2124	2950	0.8	2146	3648
4560	7029	1.0	2048	2814	1.0	2512	4215
4811	7383	1.0	1979	2691	1.2	2832	4692
5120	7783	1.0	1883	2520	1.5	3237	5263
5501	8203	1.0	1740	2271	2.0	3761	5932
5951	8499	1.0	1510	1880	3.0	4441	6619

Recalling that biomass is proportional to catch per unit of effort (Chapter 9) Fig. 5.3.1 illustrates the important conflict between the desire to maximize the total yield from a fishery, by weight or by value, and the need to give the fishermen and boat owners the necessary income. The catch per boat decreases steadily as the effort increases and may in practice become too small to make the fishery profitable, already at efforts smaller than what corresponds to the maximum on the curve for total value of the yield.

The shrimp fishery in Kuwait waters is composed of an "artisanal fishery" and an "industrial fishery". Table 5.3.3. shows the split of the total fishing mortality given in Table 5.3.1 into an artisanal component and an industrial component (from Garcia and van Zalinge, 1982). In general such a partitioning of fishing mortality into fleet components is made in the same proportions as the numbers caught.

The fishing mortality exerted by one fleet, say, fleet no. i, F(i), is

$$F(i) = F_{total} * C(i)/C_{total} \tag{5.3.1}$$

where C(i) is the number of shrimps (or fish) caught by fleet no. i, and F(total) and C(total) are the fishing mortality and the numbers caught by all fleets. F(total) may be derived from cohort analysis. The split of the predicted catch (column D in Table 5.3.1) into fleet components is obtained by:

$$C(i) = C_{total} * F(i)/F_{total} \tag{5.3.2}$$

Thus, yield and value of yield are easily separated into fleet components. Table 5.3.4 shows the split of total yield and value of yield given in Table 5.3.2 between the two fleets.

In this case the same factors, X1 = X2 are applied to the F-values of both fleets, i.e. in the exercise demonstrated in Table 5.3.4 we assume that the effort of the artisanal fleet is always the same proportion of the total effort. Fig. 5.3.2 shows the graphs corresponding to Table 5.3.4.

Table 5.3.5 and Fig. 5.3.3 show an example where the factor, X1, for the artisanal fleet is kept constant whereas the factor X2, for the industrial fleet is varied. This set up corresponds to a situation where the industrial fishery is regulated, whereas the artisanal fishery is not exposed to management measures and is assumed to remain at the same level. Note that the artisanal fleet gets a smaller share of the total catch the more effort the industrial fleet uses. This is what we expect as an increase in the effort of the industrial fleet reduces the stock so that a smaller share is left for the artisanal fleet.

It is possible to assess in a similar manner the effect of any regulatory measure for each fleet component as long as one can transfer the effort regulation into the proper fishing mortalities. For example, Garcia and van Zalinge (1982) used the Thompson and Bell model to assess the effect of a closed season.

It may also be used to assess the effect of a mesh change. In this case the selection curve for the current fishery should be estimated using one of the methods described in Chapter 6. We may express the current fishing mortality by the model (cf. Eq. 6.1.11):

$$F_{current}(t) = Fm * S_{current}(t)$$

where Fm is the maximum fishing mortality, S(current) is the selection ogive for the current gear, for example (cf. Eq. 6.1.8):

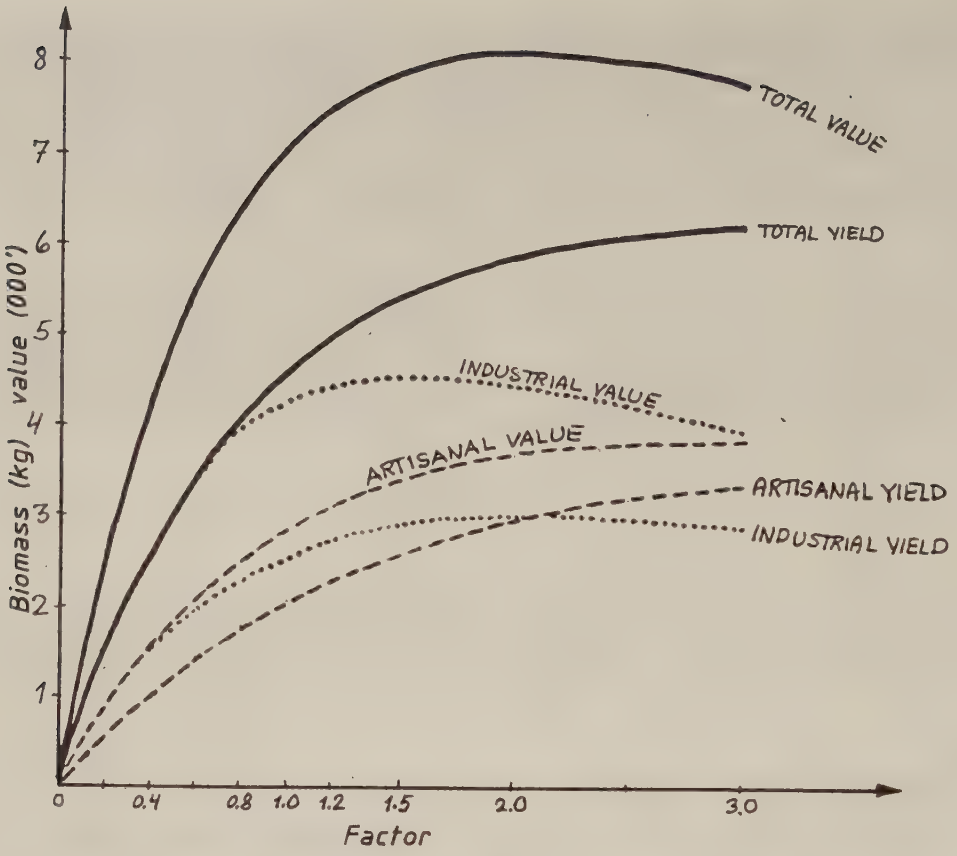


Fig. 5.3.2 Total yield and total value of yield from Fig. 5.3.1 separated into an artisanal and an industrial component

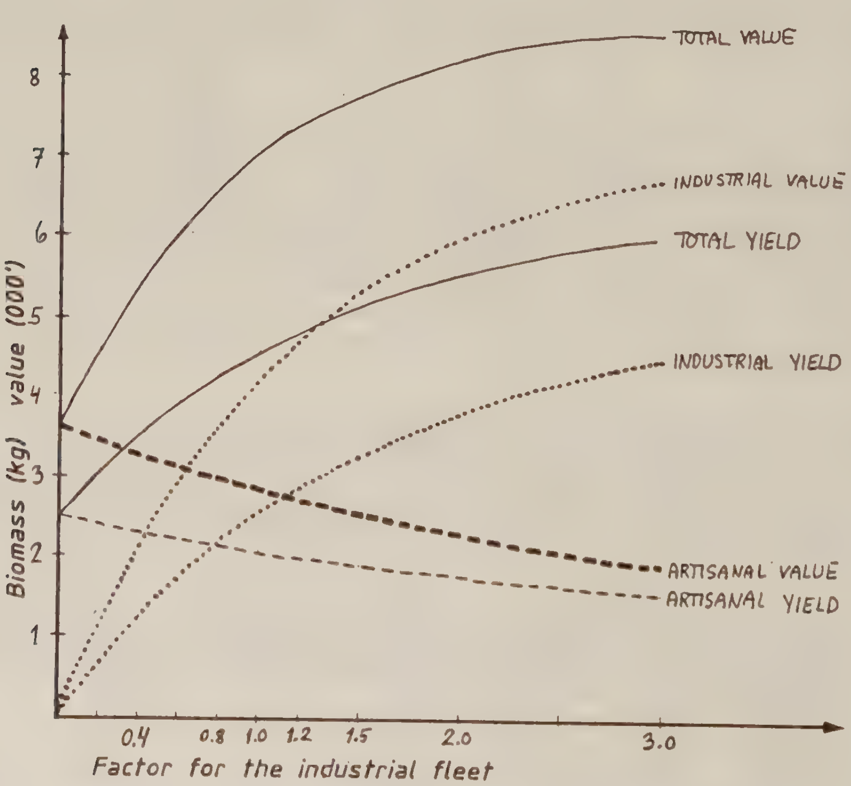


Fig. 5.3.3 Assessment of the effect of regulating the industrial fishery while the artisanal fishery is kept at a constant level

$$S_{current}(t) = 1/(1+\exp(T1-T2 * t))$$

if the gear has the trawl type of selection ogive. The parameters T1 and T2 are defined by T1 = t50% * ln3/(t75%-t50%) and T2 = T1/t50% (cf. Eqs. 6.1.9 and 6.1.10). The parameters t50% and t75% are the ages at which 50% and 75% of the fish are retained by the gear, respectively. Usually we know the lengths L50% and L75% form which the corresponding ages t50% and t75% are derived using the von Bertalanffy growth equation.

With the parameters L50% and L75% for the gear currently in use we are in a position to calculate the new selection curve under the assumption of a change of L50% and L75%. We may for example assume a 20 percent increase of L50% and L75%. From the new values of L50% and L75% we calculate new values of t50% and t75% and a new selection ogive, S(new), using Eq. 6.1.11 and finally we get the new fishing mortalities:

$$F_{new}(t) = Fm * S_{new}(t)$$

These F's are then used as input to the Thompson and Bell model, and the results for the alternative F-patterns, F(current) and F(new) can be compared (Hoydal, Rørvik and Sparre 1980 and 1982). This method is a generalization of the methods suggested by Gulland (1961), Jones (1961) and Kimura (1977).

The above described application of the Thompson and Bell model (including the mesh assessment) is essentially the method applied today to predict catches and to set catch quotas in the ICES area (Northeast Atlantic) and in many other places.

(See **Exercise(s)** in Chapter 17).

5.4 Jones' length based cohort analysis

With this section we are back again to the historic models and deal with the situation when only length composition data for the total fishery are available for one year (or the average length composition for a sequence of years). The approach is basically the same as for the length converted catch curve (cf. Section 4.4.5), i.e. we convert ages into lengths. The name "length based cohort analysis" is somewhat misleading, as we are not dealing with real cohorts in the present analysis. The real cohrot is replaced by the "pseudo-cohort" we obtain by assuming a constant parameter system (cf. Section 4.4.1). Thus, we assume that the picture presented by all size (or age) classes caught during one year reflects that of a cohort during its lifespan. We shall come back to this aspect later on.

Table 5.4.1 shows a data set for the hake fishery off Senegal (CECAF, 1978), which can be used as input for length based cohrot analysis. As was the case for the catch curve analysis (cf. Section 4.4.5) the length groups can be converted into age intervals by the inverse von Bertalanffy equation (Eq. 3.3.3.3 and Eq. 4.4.5.1 respectively):

$$t(L1) = t_o - \frac{1}{K} \ln \left[1 - \frac{L1}{L_\infty}\right] \qquad (5.4.1)$$

$$\Delta t = t(L2) - t(L1) = \frac{1}{K} \ln \left[\frac{L\infty - L1}{L\infty - L2}\right] \qquad (5.4.2)$$

For the hake off Senegal (Table 5.4.1) the von Bertalanffy parameters have been estimated as:

$$K = 0.1 \text{ per year} \quad \text{and} \quad L_\infty = 130 \text{ cm}$$

Table 5.4.1 Length composition of the hake (*Merluccius merluccius*) off Senegal (from CECAF, 1978)

length group cm L1-L2	number caught (000') C(L1,L2)
6-12	1823
12-18	14463
18-24	25227
24-30	8134
30-36	3889
36-42	2959
42-48	1871
48-54	653
54-60	322
60-66	228
66-72	181
72-78	96
78-84	16
84-∞	46

Table 5.4.2 Length groups of the hake (*Merluccius merluccius*) off Senegal, converted into age intervals, and the factors X(L1,L2). K = 0.1 per year, L_∞ = 130 cm, M = 0.28 per year

length group cm L1-L2	relative age *) t(L1)	Δt	$\exp\left(\frac{M\Delta t}{2}\right)$ X(L1,L2)	number caught (000') C(L1,L2)
6-12	0.473	0.496	1.0719	1823
12-18	0.968	0.522	1.0758	14463
18-24	1.490	0.551	1.0801	25227
24-30	2.041	0.583	1.0850	8134
30-36	2.624	0.619	1.0905	3889
36-42	3.242	0.660	1.0967	2959
42-48	3.902	0.706	1.1039	1871
48-54	4.608	0.760	1.1122	653
54-60	5.368	0.822	1.1220	322
60-66	6.190	0.890	1.1337	228
66-72	7.087	0.984	1.1478	181
72-78	8.071	1.092	1.1652	96
78-84	9.163	1.226	1.1873	16
84-∞	10.389	-	-	46

*) from Eq. 5.4.1 with t_o = 0.

Putting $t_o = 0$ (Section 4.4.5) and applying Eqs. 5.4.1 and 5.4.2 gives the results shown in Table 5.4.2, columns two and three.

To convert the cohort analysis equation (Eq. 5.2.3) into a length based version, only the term $\exp(M\,\Delta t/2)$ has to be changed.

$$\exp(M\,\Delta t/2) = \exp\left[\frac{M}{2}\,\frac{1}{K}\,\ln\left(\frac{L\infty - L1}{L\infty - L2}\right)\right] =$$

$$= \exp\left[\ln\left(\frac{L\infty - L1}{L\infty - L2}\right)^{M/2K}\right] = \left[\frac{L\infty - L1}{L\infty - L2}\right]^{M/2K}$$

Column four, $X(L1,L2)$, of Table 5.4.2 gives the values of $\exp(M\,\Delta t/2)$ for hake off Senegal with $M = 0.28$ per year and $K = 0.1$ per year, i.e. $M/2K = 0.28/(2 * 0.1) = 1.4$ (CECAF, 1978). We introduce the symbols:

$N(L1) = N(t(L1))$ = the number of fish that attain length L1
= the number of fish that attain age $t(L1)$

$N(L2) = N(t(L1) + \Delta t)$ = the number of fish that attain length L2
= the number of fish that attain age $t(L2)$ ($= t(L1) + \Delta t$)

$C(L1,L2) = C(t,t + \Delta t)$ = the number of fish caught of lengths between L1 and L2
= the number of fish caught of ages between $t(L1)$ and $t(L2)$.

$$X(L1,L2) = \left[\frac{L\infty - L1}{L_\infty - L2}\right]^{M/2K} \qquad (5.4.3)$$

$$= \left[\begin{array}{l}\text{the fraction of } N(L1) \text{ which survive}\\ \text{natural deaths during the time period}\\ \text{from } t(L1) \text{ to } t(L1) + \Delta t/2\end{array}\right]^{-1}$$

Eq. 5.2.3 can be written:

$$N(L1) = [N(L2)\,X(L1,L2) + C(L1,L2)]\,X(L1,L2) \qquad (5.4.4)$$

The calculation procedure of Eq. 5.4.4 is similar to that of the age based cohort analysis (Eq. 5.2.1). We start with the last group and use the catch equation. In the case of the hake off Senegal the last group is the hake of length 84 cm and longer:

$$C(84,\infty) = N(84)\,\frac{F}{Z}\,[1 - \exp(-Z\,\Delta t)]$$

In this case Δt refers to all fish longer than 84 cm, so that Δt is very large. Theoretically, the age corresponding to L_∞ is ∞, so theoretically $\Delta t = \infty$, and therefore:

$$\exp(-Z * \infty) = 0$$

Thus, approximately:

$$C(84,\infty) = N(84)\,\frac{F}{Z} \qquad (5.4.5)$$

or

$$N(84) = \frac{C(84,\infty)}{F/Z} = \frac{46}{F/Z} \quad \text{(cf. Table 5.4.2)}$$

In the length converted cohort analysis we usually make the initial guess of the exploitation rate, F/Z, rather than the terminal F. If F/Z for the last length group is assumed to take the value 0.5 then the number of hake attaining length 84 cm becomes:

$$N(84) = 46/0.5 = 92$$

Inserting into Eq. 5.4.4 gives N(78), the number of hake attaining length 78 cm:

$$N(78) = (92 * 1.1873 + 16) * 1.1873 = 148.7$$

as $X(L1,L2) = 1.1873$ (cf. Table 5.4.2).

Continuing to move backwards in length (and thus in time) the following stock numbers are calculated:

$$N(72) = (148.7 * 1.1652 + 96) * 1.1652 = 313.7$$

$$N(66) = (313.7 * 1.1478 + 181) * 1.1478 = 621.0$$

.... etc. ....

The stock numbers derived this way are given for all length groups in Table 5.4.3.

Table 5.4.3 The calculation procedure of Jones' length cohort analysis illustrated by the hake (Merluccius merluccius) off Senegal. K = 0.1 per year, L_∞ = 130 cm, M = 0.28 per year

length group cm	a)	number caught (000')	number of survivors b)	exploitation rate c)	fishing mort. d)	total mort. e)
L1-L2	X(L1,L2)	C(L1,L2)	N(L1)	F/Z	F	Z
6-12	1.0719	1823	98919.3	0.1255	0.04	0.32
12-18	1.0758	14463	84392.7	0.5805	0.39	0.67
18-24	1.0801	25227	59475.8	0.7920	1.07	1.35
24-30	1.0850	8134	27623.0	0.6979	0.65	0.93
30-36	1.0905	3889	15967.8	0.6369	0.49	0.77
36-42	1.0967	2959	9861.5	0.6785	0.59	0.87
42-48	1.1039	1871	5500.5	0.6977	0.65	0.93
48-54	1.1122	653	2818.8	0.5792	0.39	0.67
54-60	1.1220	322	1691.5	0.5072	0.29	0.57
60-66	1.1337	228	1056.6	0.5234	0.31	0.59
66-72	1.1478	181	621.0	0.5890	0.40	0.68
72-78	1.1652	96	313.7	0.5817	0.39	0.67
78-84	1.1873	16	148.7	0.2823	0.11	0.39
84-∞	-	46	92.0 f)	0.5000	0.28	0.56

a) $X(L1,L2) = \left[\frac{L\infty - L1}{L\infty - L2} \right]^{M/2K} = \left[\frac{130 - L1}{130 - L2} \right]^{1.4}$

b) $N(L1) = [N(L2)\ X(L1,L2) + C(L1,L2)]\ X(L1,L2)$

c) $F/Z = C(L1,L2)/[N(L1)-N(L2)]$

d) $F = M\ (F/Z)/(1 - F/Z)$

e) $Z = F + M$

f) $N(84) = C(84,\infty)/(F/Z) = 46/0.5 = 92$

To estimate F we could use Eq. 5.2.4, but it is more convenient to calculate F by:

$$F = M \frac{F/Z}{1 - F/Z} \qquad (5.4.6)$$

where the exploitation rate F/Z is derived from:

$$F/Z = \frac{C(L1,L2)}{N(L1) - N(L2)} \qquad (5.4.7)$$

For example for length group 72-78 cm we get (cf. Table 5.4.3):

$$F/Z = \frac{C(72,78)}{N(72) - N(78)} = \frac{96}{313.7 - 148.7} = 0.5817$$

$$F = M \frac{F/Z}{1 - F/Z} = 0.28 \frac{0.5817}{1 - 0.5817} = 0.39$$

The complete calculation procedure for length cohort analysis is given in Table 5.4.3.

We now want to calculate the mean number of fish in the sea and their biomass. Summing the column of N(L1) values in Table 5.4.3 would not give the right number because changing the class interval would give a different sum: the N(L1) values are simply the number alive at any length L1. The procedure is to find the mean number in each class interval and weight it by the time, Δt, spent) in that class interval. It is the same problem dealt with by Eqs. 4.2.6 and 4.2.9. Some manipulation of either of these leads to the result:

$$NMEAN(L1,L2) * \Delta t = (N(L1) - N(L2))/Z \qquad (5.4.8)$$

which is the ANNUAL mean number in each length class. The total mean number of fish in the sea of lengths above the first L1 (here 6 cm) becomes:

$$\sum_i NMEAN(Li,Li+1) * \Delta t \qquad (5.4.9)$$

Correspondingly, we find the annual mean biomass in each length group by multiplying by the mean weight, W(L1,L2), in the length group:

$$BIOM\ (L1,L2) * \Delta t = NMEAN(L1,L2) * \Delta t * W(L1,L2) \qquad (5.4.10)$$

The body weight is calculated from

$$W(L1,L2) = q * (L1^b + L2^b)/2 \qquad (5.4.11)$$

in which q and b are the constants of the length/weight relationship described in Section 2.6. The body weight of the last group may be calculated as $W(L1,L_\infty)$.

The sum

$$\sum_i BIOM(Li,Li+1) * \Delta t \qquad (5.4.12)$$

in an estimate of the average biomass during the lifespan of a cohort, or of all cohorts during a year, and is independent of the length class interval.

The body weight may also be used to estimate the yield, i.e. the weight of the catch. The weight of the catch belonging to length group i becomes

$$\sum_{i} C(Li,Li+1) * W(Li,Li+1) \qquad (5.4.13)$$

Table 5.4.4 shows the calculation of the yield and the average biomass during a year.

The length cohort analysis defined by Table 5.3.3 is called "Jones' length cohort analysis". (Jones, 1976, and Jones and van Zalinge, 1981, reviewed in Jones, 1984 and Pauly, 1984). As already mentioned, the method is usually applied to "pseudo cohorts", i.e. we assume a constant parameter (equilibrium) system. In order to simulate an equilibrium condition it is essential that the data pertain to a relatively long time period (say a year or several years), preferably a number of full years.

A length frequency sample collected during a relatively short time period is not applicable. The September sample of shrimps shown in Fig. 3.4.2.5 consisting of only one cohort is such an example. The method assumes that a larger specimen is also older, but in this case we assumed that all shrimps, irrespective of length, are of the same age. The descending slope of that sample has something to do with the variation in individual growth rates - it is not related to mortality.

Table 5.4.4 The calculation procedure for yield and average biomass in Jones' length based cohort analysis illustrated by the hake (*Merluccius merluccius*) off Senegal. $q = 0.00001$ kg/cm^3, $b = 3$, $K = 0.1$ per year, $L_\infty = 130$ cm, $M = 0.28$ per year

length group cm	number caught (000')	number of survivors (000')	total mort. rate	body weight kg	mean N * Δt (000')	mean biomass * Δt tonnes	yield tonnes
L1-L2	C	N(L1)	Z	W a)	b)	c)	d)
6-12	1823	98919.3	0.32	0.0097	45369	441.0	17.7
12-18	14463	84392.7	0.67	0.0378	37335	1411.3	546.7
18-24	25227	59475.8	1.35	0.0983	23664	2325.7	2479.3
24-30	8134	27623.0	0.93	0.204	12575	2566.9	1660.3
30-36	3889	15967.8	0.77	0.368	7919	2916.4	1432.2
36-42	2959	9861.5	0.87	0.605	5007	3022.8	1786.4
42-48	1871	5500.5	0.93	0.923	2895	2673.6	1727.9
48-54	653	2818.8	0.67	1.34	1694	2270.5	875.2
54-60	322	1691.5	0.57	1.87	1117	2086.6	601.3
60-66	228	1056.6	0.59	2.52	741	1866.4	574.0
66-72	181	621.0	0.68	3.30	451.1	1490.3	598.0
72-78	96	313.7	0.67	4.24	246.5	1045.1	407.0
78-84	16	148.7	0.39	5.34	144.9	775.2	85.4
84- ∞	46	92.0	0.56	6.59	164.3	1085.7	304.0
					Total	25977.5	13095.1

a) $W(L1,L2) = q * (L1^b + L2^b)/2$
b) $(N(L1)-N(L2))/Z = (\text{mean number}) * \Delta t)$
c) $W(L1,L2) * (N(L1)-N(L2))/Z$
d) $W(L1,L2) * C(L1,L2)$

It is possible to apply Jones' method to a real cohort, but that implies that we are able to follow a cohort through time, i.e. that we know its age. If that is the case, however, we might as well use ordinary age based cohort analysis, as this does not present problems in connection with the conversion of length into age.

As Jones' length based cohort analysis is based on Pope's age based cohort analysis (Section 5.2) it has the same limitations. The approximation to VPA is valid for values of F Δt up to 1.2 and M Δt up to 0.3 (Pope, 1972). A version of length based VPA (i.e. a method not requiring limitations of the values of F Δt and M Δt) was presented in Sparre (1979).

The program "LCOHOR" in the LFSA package of microcomputer programs (Sparre, 1987) can execute Jones' length based cohort analysis as described above. Also the "COMPLEAT ELEFAN" package (Gayanilo, Soriano and Pauly, 1988) contains routines for a length based analysis similar to LCOHOR as well as a length based method similar to VPA (i.e. it does not work with pseudo-cohorts).

(See **Exercise(s)** in Chapter 17).

5.5 A predictive length based model (Thompson and Bell)

The "length based Thompson and Bell model" is the "forwards" version of the length based cohort analysis. It is used to predict catches and stock sizes under given assumptions on future exploitation levels and mesh sizes. The analysis takes the fishing mortalities by length group as input and calculates the numbers caught as well as the stock numbers. It further requires the number of fish in the smallest length group and the natural mortality as input. The natural mortality must be the same as for the cohort analysis which provides also an estimate of the numbers in the first length group.

To turn Eqs. 5.4.4 and 5.4.7 into a "forward version" Eq. 5.4.7 is rewritten as follows:

$$C(L1,L2) = (F/Z) * (N(L1) - N(L2)) \qquad (5.5.1)$$

which inserted into Eq. 5.4.4 gives:

$$N(L1) = (N(L2)\ X(L1,L2) + (F/Z)\ (N(L1) - N(L2)))\ X(L1,L2)$$

Solving this equation with respect to N(L2) gives:

$$N(L2) = N(L1) * (1/X(L1,L2) - F/Z) / (X(L1,L2) - F/Z) \qquad (5.5.2)$$

in which

$$X(L1,L2) = ((L_\infty - L1)/(L_\infty - L2))^{M/2K} \text{ as before.}$$

In its simplest form, the analysis uses the F-array estimated in cohort analysis as the reference F-array and assesses the effect of raising (reducing) all F's by a certain factor, as was the case for the analysis based on age composition data (see Section 5.3). We start with the number in the smallest length group, say, N(L1), calculated by cohort analysis. The number of survivors attaining length L2 is:

$$N(L2) = N(L1) * (1/X(L1,L2)-E(L1,L2))/(X(L1,L2) - E(L1,L2)) \quad \text{where}$$

$$E(L1,L2) = F(L1,L2)/Z(L1,L2)$$

The number caught of lengths between L1 and L2 is:

$$C(L1,L2) = F(L1,L2) * (N(L1)-N(L2))/Z(L1,L2)$$

In the above example the F-value estimated for fish of lengths between L1 and L2 from cohort analysis has been applied. The general case is dealt with later on. Hereby the calculation for the first length group is completed, and the second length group, from L2 to L3, is treated in the same way:

$$N(L3) = N(L2) * (1/X(L2,L3)-E(L2,L3))/(X(L2,L3) - E(L2,L3)) \quad \text{where}$$

$$E(L2,L3) = F(L2,L3)/Z(L2,L3)$$

$$C(L2,L3) = F(L2,L3) * (N(L2)-N(L3))/Z(L2,L3)$$

The general step (no. i) thus becomes:

$$N(Li+1) = N(Li) * (1/X(Li,Li+1)-E(Li,Li+1))/(X(Li,Li+1)-E(Li,Li+1))$$

where

$$E(Li,Li+1) = F(Li,Li+1)/Z(Li,Li+1)$$
$$C(Li,Li+1) = F(Li,Li+1) * (N(Li)-N(Li+1))/Z(Li,Li+1)$$

In the more general case where all F-values are raised (or reduced) by the factor XX the general step becomes:

$$N(Li+1) = N(Li) * (1/X(Li,Li+1)-E(Li,Li+1))/(X(Li,Li+1)-E(Li,Li+1)) \tag{5.5.3}$$

where

$$E(Li,Li+1) = XX * F(Li,Li+1)/Z(Li,Li+1)$$
$$Z(Li,Li+1) = XX * F(Li,Li+1) + M$$

$$C(Li,Li+1) = XX * F(Li,Li+1) * (N(Li)-N(Li+1))/Z(Li,Li+1) \tag{5.5.4}$$

The yield (catch in weight) in length group i is:

$$YIELD(Li,Li+1) = C(Li,Li+1) * W(Li,Li+1)$$

where W(Li,Li+1) is the mean weight of fish of lengths between Li and Li+1. W(Li,Li+1) may be calculated as before from:

$$W(Li,Li+1) = q * (Li^{b} + Li+1^{b})/2$$

where q and b are the parameters in the length/weight relationship.

The value of the yield is given by:

$$VALUE(Li,Li+1) = YIELD(Li,Li+1) * PRICE(Li,Li+1) \quad \text{where}$$

PRICE(Li,Li+1) is the price per kg of fish between lengths Li and Li+1

The mean number of survivors times $\underline{/\backslash}t$, where $\underline{/\backslash}t$ is the time it takes to grow from length Li to Li+1 in length group i, is given by Eq. 5.4.8:

$$NMEAN(Li,Li+1) * \underline{/\backslash}t = (N(Li) - N(Li+1))/Z(Li,Li+1)$$

and the corresponding mean biomass times $\underline{/\backslash}t$ is

$$BIOM(Li,Li+1) * \underline{/\backslash}t = NMEAN(Li,Li+1) * \underline{/\backslash}t * W(Li,Li+1)$$

as discussed in Section 5.4. The sum:

$$\sum_{i} \mathrm{BIOM}(Li,Li+1) * \Delta t$$

is an estimate of the average biomass during the lifespan of a cohort, or all cohorts during a year. Comparing with the age based method, Section 5.3, it is seen that it was not necessary there to multiply each biomass by Δt because this was constant and equal to 1/12 of a year, or one month. It was easier to divide the sum of biomasses by 12.

The length based Thompson and Bell analysis is derived from Jones' length based cohort analysis (Section 5.4) which in turn is based on Pope's cohort analysis (Section 5.2). The length based Thompson and Bell method has therefore the same limitations as Pope's cohort analysis.

The approximation to VPA in forecast mode is valid for values of $F\ \Delta t$ up to 1.2 and $M\ \Delta t$ up to 0.3 (Pope, 1972). If the F's are high, nonsense results will come out of the analysis, such as negative stock numbers. Smaller length groups and hence, smaller Δt values, are required.

As an example we use the data from Table 5.4.3 (*Merluccius merluccius*) caught off Senegal. The following input parameters are used (compare Section 5.4):

L_∞ = 130 cm, K = 0.1 per year, M = 0.28 per year, q = 0.00001 kg/cm^3, b = 3, N(first length group) = 98919.3

Using the F-values from Table 5.4.3 and the body weights derived from the length/weight relationship, and some (in this case arbitrarily selected) prices per kg for hake, the input may be summarized as in Table 5.5.1. Using Eq. 5.5.3 with XX = 1 we get

$$\begin{aligned} N(12) &= N(6) * (1/X(6,12)-E(6,12))/(X(6,12)-E(6,12)) \\ &= 98919.3 * (1/1.0719 - .125)/(1.0719 - .125) \\ &= 84392.7 \quad \text{where} \\ E(6,12) &= .04/(.28+.04) = .125 \\ X(6,12) &= 1.0719 \text{ (taken from Table 5.4.3)} \end{aligned}$$

Actually the calculation above gives N(12) = 84400.8. The correct figure is obtained by using 8 significant digits in all calculations.

Table 5.5.1 Input data for length based Thompson and Bell analysis

length (Li)	F(Li,Li+1)	W(Li,Li+1)	PRICE(Li,Li+1)
6	0.04	0.0097	1.0
12	0.39	0.0378	1.0
18	1.07	0.0983	1.0
24	0.65	0.204	1.5
30	0.49	0.368	1.5
36	0.59	0.605	2.0
42	0.65	0.923	2.0
48	0.39	1.34	2.5
54	0.29	1.87	2.5
60	0.31	2.52	2.5
66	0.40	3.30	3.0
72	0.39	4.24	3.0
78	0.11	5.34	3.0
84	0.28	6.59	3.0

In the following calculations the values corresponding to Table 5.4.3 (calculated with 8 significant digits) are given as results to show that we are now in a process of reconstructing Table 5.4.3 in our prediction. The number caught is:

$$\begin{aligned} C(6,12) &= F(6,12) * (N(6) - N(12))/Z(6,12) \\ &= .04 * (98919.3 - 84392.7)/.32 \\ &= 1823 \end{aligned}$$

The next step is

$$\begin{aligned} N(18) &= N(12) * (1/X(12,18)-E(12,18))/(X(12,18)-E(12,18)) \\ &= 84392.7 * (1/1.0758-.582)/(1.0758-.582) \\ &= 59475.8 \quad \text{where} \\ E(12,18) &= .39/(.28+.39) = .582 \\ X(12,18) &= 1.0719 \text{ (taken from Table 5.4.3)} \\ C(12,18) &= F(12,18) * (N(12) - N(18))/Z(12,18) \\ &= .39 * (84392.7 - 59475.8)/.67 \\ &= 14463 \end{aligned}$$

Continuing this way, the final result would be the values given in Table 5.4.3. Adding the calculations of yield, biomass and value the results become as shown in Table 5.5.2 (compare Table 5.4.4).

As the last length group is a plus group it might have been given a special treatment (compare Section 5.4.3):

$$\begin{aligned} C(84,\infty) &= N(84) * F(84,\infty/Z(84,\infty) \\ &= 92 * .28/(.28 + .28) \\ &= 46 \end{aligned}$$

$$NMEAN(84,\infty) = N(84)/Z(84,\infty) = 92/.56 = 164.3$$

$$BIOM(84,\infty) = NMEAN(84,\infty) * W(84,\infty) = 164.3 * 6.59 = 1082.6$$

Table 5.5.3 shows the results corresponding to Table 5.5.2 but with the F-factor 2.0, i.e. the prediction of catch under the assumption of a doubling of the fishing effort. In this case the effect of doubling the effort would be a dramatic fall in the yield and the value of the yield.

Table 5.5.2 Output from length based Thompson and Bell analysis, using the F-factor 1.0. Weights are in tonnes

length	N	mean biomass	C	yield	F	value
6	98919.3	441.0	1823	17.7	0.04	17.7
12	84392.7	1411.3	14463	546.7	0.39	546.7
18	59475.8	2325.7	25277	2479.3	1.07	2479.3
24	27623.0	2566.9	8143	1660.3	0.65	2490.5
30	15967.8	2916.4	3889	1432.2	0.49	2148.4
36	9861.5	3022.8	2959	1786.4	0.59	3572.8
42	5500.5	2673.6	1871	1727.9	0.65	3455.4
48	2818.8	2270.5	653	875.2	0.39	2188.0
54	1691.5	2086.6	322	601.3	0.29	1503.2
60	1056.6	1866.4	228	574.0	0.31	1435.0
66	621.0	1490.3	181	598.0	0.40	1793.9
72	313.7	1045.1	96	407.0	0.39	1220.8
78	148.7	775.2	16	85.4	0.11	256.1
84	92.0	1085.7	46	304.0	0.28	912.0
Total		25977.5	59908	13095.1		24019.8

Table 5.5.4 shows the summary results for 16 different F-factors. Each row is based on calculations like those illustrated by Tables 5.5.2 and 5.5.3. The total yield, mean biomass and value given in the last row of these tables can be found also in Table 5.5.4. The two last rows show the maximum yield (MSY) and the maximum value (MSE) together with the corresponding F-factor and stock biomass. When kg prices differ from one length group to another the F-factor giving MSY usually differs from the F-factor giving MSE.

Table 5.5.3 Output from length based Thomson and Bell analysis, using the F-factor 2.0. (Compare Table 5.5.2)

length	N	mean biomass	C	yield	F	value
6	98919.3	436.8	3611.6	35.1	0.08	35.1
12	82724.1	1270.5	26041.2	984.4	0.77	984.4
18	47271.6	1514.8	32863.1	3229.8	2.13	3229.8
24	10092.9	813.3	5154.2	1052.1	1.29	1578.1
30	3823.1	619.6	1652.3	608.5	0.98	912.8
36	1699.7	450.3	881.6	532.3	1.18	1064.5
42	609.2	251.3	351.7	324.7	1.29	649.5
48	181.3	130.2	74.9	100.4	0.77	250.9
54	79.3	88.8	27.4	51.2	0.58	128.0
60	38.5	61.0	14.9	37.5	0.62	93.9
66	16.8	34.9	8.5	28.0	0.80	84.1
72	5.4	15.4	2.8	12.0	0.78	35.9
78	1.5	7.6	0.3	1.7	0.22	5.0
84	0.8	6.5	0.5	3.6	0.56	10.9
Total		5701.0	70685.0	7001.3		9062.8

Table 5.5.4 Yield table derived by the length based Thompson and Bell analysis. MSY = Maximum Sustainable Yield. MSE = Maximum Sustainable Economic Yield (value)

X	yield	mean biomass	value
0.0	0	450340	0
0.2	16566	224890	40408
0.4	19117	119846	43768
0.6	17628	67893	37630
0.8	15296	40807	30302
1.0	13095	25978	24020
1.2	11262	17473	19128
1.4	9797	12374	15444
1.6	8641	9183	12691
1.8	7728	7105	10626
2.0	7001	5701	9063
2.2	6417	4719	7864
2.4	5941	4011	6934
2.6	5549	3486	6201
2.8	5223	3088	5618
3.0	4949	2780	5147

MSY = 19132 for factor = .385 Biom.msy = 122571
MSE = 44453 for factor = .335 Biom.mse = 142706

Tables 5.5.2-5.5.4 were calculated by the program "MIXFISH" in the LFSA package (Sparre, 1987). This program calculates MSY and MSE by using an iterative technique.

In case of economic interaction (several fleets) exploiting one resource the catches predicted by the length based Thompson and Bell analysis can be partitioned exactly as done in Section 5.3 (cf. Tables 5.3.3-5.3.5).

The effect of a change of mesh size can be assessed by changing the array of F-values (the fishing pattern) accordingly. The models for gear selection introduced in Chapter 6 may be applied to change the fishing pattern corresponding to a certain change of selection. That is, we use Eq. 6.1.11:

$$F_{current}(L) = Fm * S_{current}(L)$$

where Fm is the maximum fishing mortality and S(current) the selection ogive for the current gear, for example (cf. Eq. 6.1.1):

$$S_{current}(L) = 1/(1+\exp(S1-S2 * t)$$

if the gear has the trawl type of selection ogive. The parameters S1 and S2 are defined by (cf. Eqs. 6.1.4 and 6):

$$S1 = L50\% * \ln 3/(L75\%-L50\%) \quad \text{and} \quad S2 = S1/L50\%$$

The parameters L50% and L75% are the lengths at which 50% and 75% of the fish are retained by the gear, respectively. With the parameters L50% and L75% for the gear currently in use we are in a position to calculate the new selection curve and thereby the new fishing pattern under the assumption of a change of L50% and L75% as was done for the age based model described in the preceding section.

The program "MIXFISH" in the LFSA package on microcomputer programs (Sparre, 1987) contains an option for mesh assessment corresponding to the above described procedure. This program produces an output table showing the total yield for various combinations of effort and L50%, i.e. a table of the form:

		Relative Effort				
		- 20 %	- 10 %	no change	+ 10 %	+ 20 %
Relative value of L50%	- 30 %					
	- 15 %					
	no change			YIELD		
	+ 15 %					
	+ 30 %					

MIXFISH assumes L75% to be proportional to L50%. MIXFISH allows you to test any combination of L50% and effort, and thereby enables you to determine the optimum combination of L50% and effort. The middle cell (marked "YIELD") corresponds to the present fishing regime.

The assumption behind the length based Thompson and Bell analysis and behind Jones' length based cohort analysis is that the stock remains in a steady state, with all parameters (e.g. recruitment) remaining constant. Thus, we obtain a prediction of the average long term catches. Deviations from the predicted catches are to be expected in individual years.

In his Appendix 4 Jones (1984), based on Jones (1961), suggests a slightly different approach to the one given above. Instead of inverting the length based cohort analysis he suggests to use the formula for the number caught from length group j (using his notation):

$$C(j) = R\ r(j)\ (F1\ dt)j\ \exp(- M(j) - \Sigma\ (F1\ dt)j) \quad \text{where}$$

$M(j)$	is the cumulative natural mortality up to the middle of length group j.
R	is the number of recruits to the smallest length group.
$r(j)$	is the proportion of fish in length group j that are retained on board. In the derivation of the length based Thompson and Bell model, this quantity is tacitly assumed to be 1, but the model can easily be changed to account for discarding.
$(F1\ dt)$	is the product of fishing mortality and dt, the time required to grow through the length class, for length group j.
$\Sigma\ (F1\ dt)j)$	is the cumulative value up to the middle of length group j.

Essentially, this model (with a slight extension) can be used for the same purposes as the length based Thompson and Bell model. The two models may give slightly different results, because they are both approximations to the exact exponential decay model. The only advantage of the length based Thompson and Bell method compared to Jones' approach is that the first one is consistent with the length based cohort analysis, from which both models take their input.

(See **Exercise(s)** in Chapter 17).

6 GEAR SELECTIVITY

So far we have dealt mostly with that part of the age (or length) ranges representing the part of the stock under full exploitation. Obviously, the observed numbers caught in the only partly exploited age (or length) groups also provide important knowledge of the exploitation of the fish stock.

In Section 4.5 the concept of L50% (length at first capture) was introduced. This chapter deals with a mathematical model for the selection ogive (cf. Fig. 4.5.3.1) and the estimation of selection parameters. As already mentioned the optimum exploitation strategy is determined by the age at first capture and the overall level of fishing mortality. Therefore, it is important to have an estimate of L50% when assessing a fishery. The actual use of length (or age) at first capture will be described in Chapter 8, whereas this section deals only with the definition and estimation of the relevant parameters.

Some gears have a descending right hand slope in their selection curve, i.e. also the largest fish have less probability of being caught. This applies e.g. to gill nets (cf. Section 6.4) and leads to a bell-shaped selection curve. Most of this chapter however, is devoted to problems relating to the trawl type of selection. Discussions of selectivity can be found in e.g. Pope et al. (1975), Gulland (1983) and Pauly (1984).

6.1 Trawl selection

Fig. 6.1.0.1 shows a selection curve with a 50% retention point of 15 cm. If all fish above 15 cm were retained and all below 15 cm escaped, the curve would be the fat vertical line. In practice some fish above the 50% retention length escape and some below are retained. We therefore find an S-shaped curve whose shape can be further described by the position of the 25% and 75% retention points (see the figure). Such selection curves can be estimated by specially designed experiments described in Section 6.1.3. However, a curve of similar shape is obtained when the fish are recruiting to the area fished, even when a very small mesh through which no fish escapes is used. This is called the "recruitment curve". In practice we often find a curve which combines the effect of recruitment and gear selection. This is called a "resultant curve". The probability that a fish is retained by the trawl is thus the product of two probabilities:

1. the probability that the fish is present on (has recruited to) the fishing ground

2. the probability that the fish is retained by the meshes once it has entered the trawl.

Fig. 6.1.0.2 illustrates these points. The curve R is the recruitment curve, G is the mesh selection curve and S = R * G the resultant curve.

The points with probability 0.5 (the 50% points) of the curves are indicated in the figure. At length Q (see the figure) we have A * B = C or R(Q) * G(Q) = S(Q): C = 0.62 * 0.42 = 0.26. At length X practically all fish have recruited to the fishing grounds while some are still small enough to escape through the meshes. At length Y even mesh retention is complete and no fish escapes.

When the meshes are so large that there is no overlap of the recruitment curve with the selection curve we can ignore recruitment. The resultant curve is then determined by selection only, see Fig. 6.1.0.3.

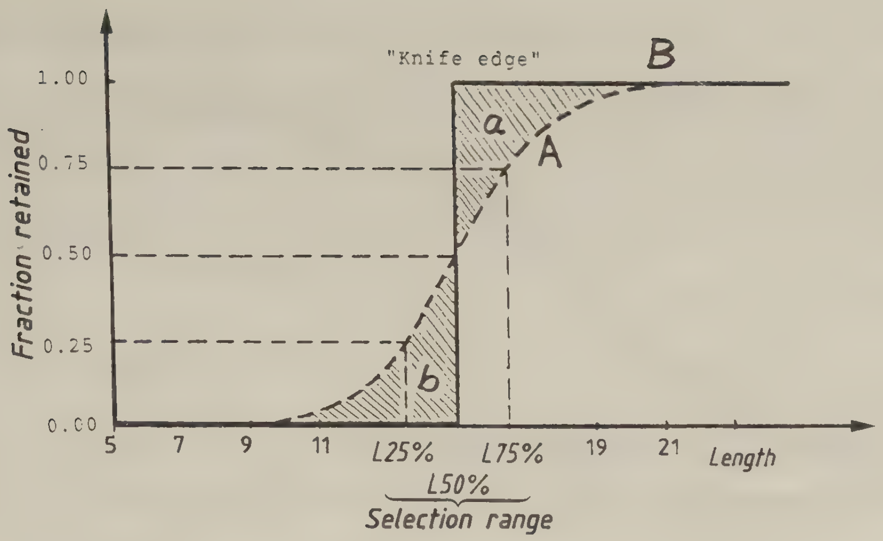


Fig. 6.1.0.1 Trawl selection curve as a function of body length. An illustration of the concepts of knife edge selection and selection range (Section 6.1)

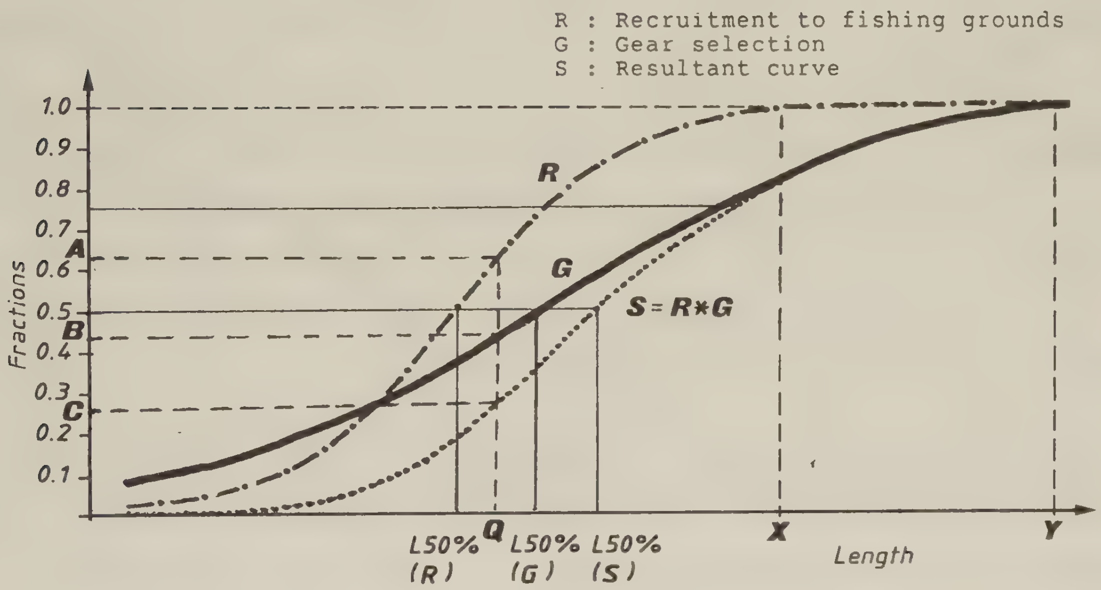


Fig. 6.1.0.2 Recruitment and selection curves and the resultant curve, see also text

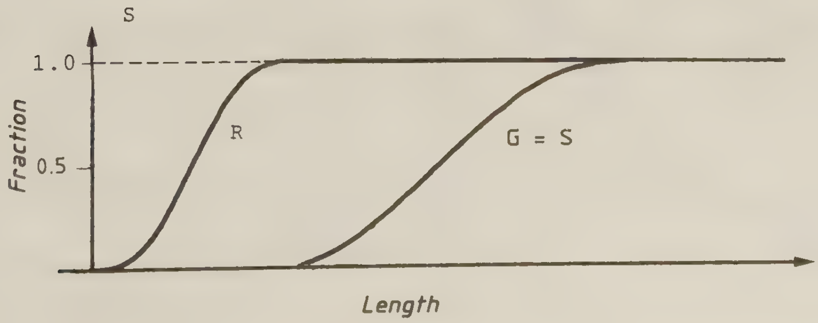


Fig. 6.1.0.3 Example of equality between the gear selection curve (G) and the resultant curve (S)

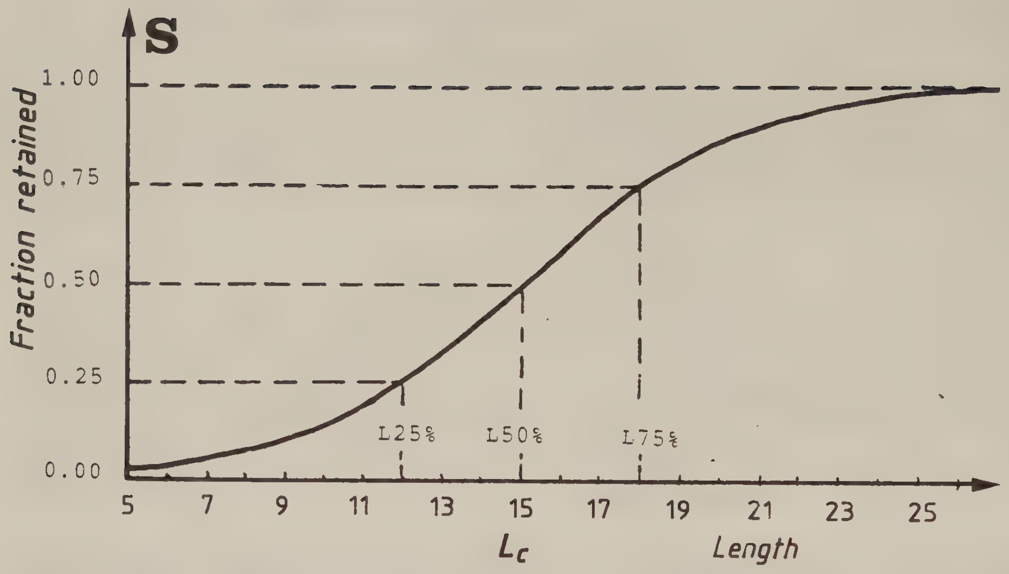


Fig. 6.1.1.1 The selection curve referred to in the text

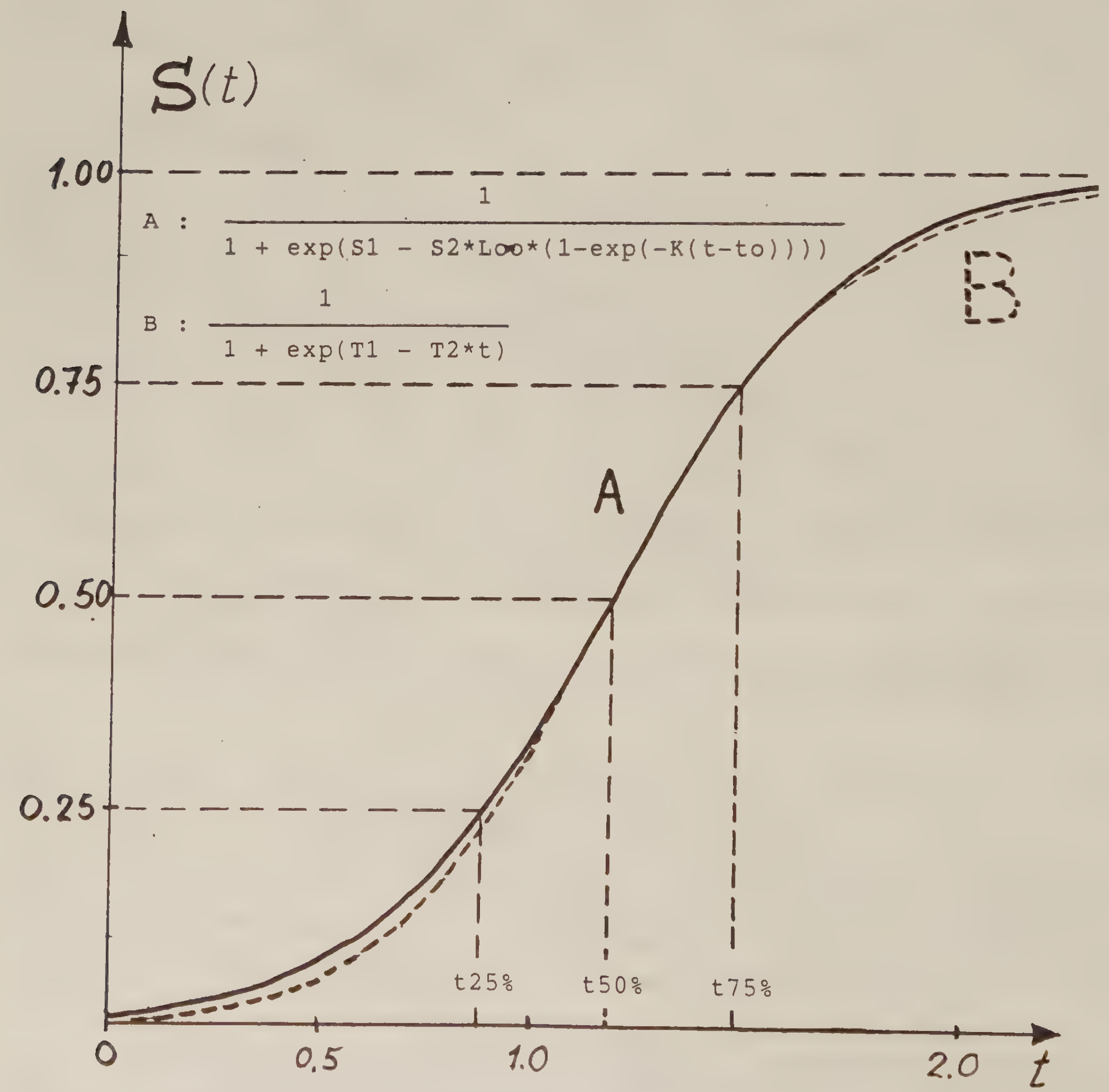


Fig. 6.1.1.2 A: Exact age transformed selection ogive (Eq. 6.1.1.7)
B: Approximated selective ogive (Eq. 6.1.1.8)

6.1.1 A mathematical model for the selection ogive

In Section 2.2 the cumulative normal distribution was introduced (cf. Figure 2.2.2) and as can be seen, this mathematical expression produces a graph with the same shape as the selection curve (Fig. 6.1.1.1). Thus, we could use the mathematical expression for the cumulative normal distribution (Eq. 2.2.3) to describe the selection ogive. (As done by Jones, 1963 and 1976). However, from a calculation point of view it is easier to work with the so-called "logistic curve":

$$S(L) = \frac{1}{1 + \exp(S1 - S2 * L)} \qquad (6.1.1.1)$$

where L is the length and S1 and S2 are constants. (Paloheimo and Cadima, 1964; Kimura, 1977 and Hoydal, Rørvik and Sparre, 1982). The curve shown in Fig. 6.1.1.1 is the graph of Eq. 6.1.1.1 with S1 = 5.4930 and S2 = 0.3662.

By applying a few algebraic manipulations it follows that there is a one to one correspondence between L50%, L75% and S1, S2:

$$L50\% = S1/S2 \qquad (6.1.1.2)$$

$$L75\% = (\ln 3 + S1)/S2 \qquad (6.1.1.3)$$

$$S1 = L50\% * \ln 3/(L75\% - L50\%) \qquad (6.1.1.4)$$

$$S2 = \ln 3/(L75\% - L50\%) = S1/L50\% \qquad (6.1.1.5)$$

In the case of Figure 6.1.1.1 we find:

$S1 = 15 * \ln 3/(18-15) = 5.4930$ and $S2 = \ln 3/(18-15) = 0.3662$

$L50\% = 5.4930/0.3662 = 15$ and

$L75\% = (5.4930 + \ln 3)/0.3662 = 18$

The main reason for choosing the mathematical expression of Eq. 6.1.1.1 is that it is easy to handle computationally. The choice is not based on biological/physical reasoning as was the case for the models of growth and mortality. Estimation procedures are dealt with in Sections 6.2 and 6.3.

The ogive defined by Eq. 6.1.1.1 is symmetrical around L50%, for example: $L25\% = (\ln 3 - S1)/S2$ so that:

$$L75\% - L50\% = \ln 3/S2 = L50\% - L25\%$$

This conclusion holds for all percentages: (50+X)% and (50-X)%.

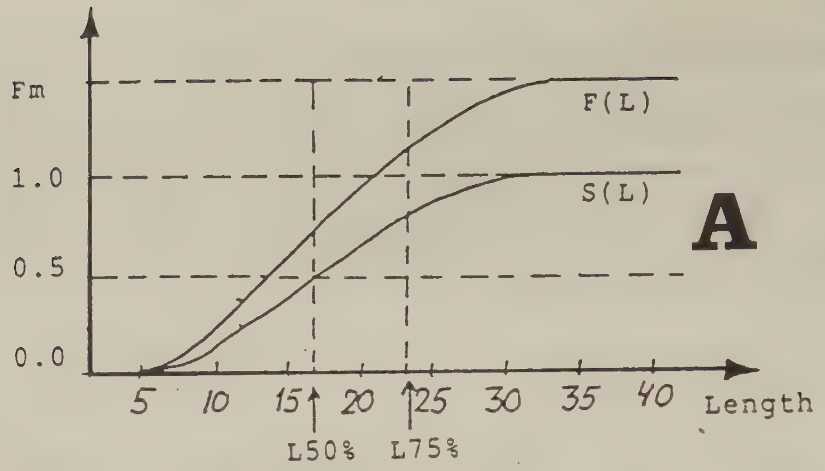
We may observe non-symmetrical selection ogives, but we often assume that a symmetrical curve is a reasonable approximation. The quantity

$$L75\% - L25\% = 2 * \ln 3/S2 \qquad (6.1.1.6)$$

is called the "selection range". In the case of Fig. 6.1.1.1 the selection range is:

$$18 - 12 = 6 \; (= 2 * \ln 3/0.3662)$$

Figure 6.1.0.1 shows two selection curves. Curve A has a selection range of 3 cm and curve B has a selection range of 0 cm. Curve B is a so-called "knife-edge" selection curve (Beverton and Holt, 1957). Knife-edge selection should be considered a hypothetical model which will never describe a



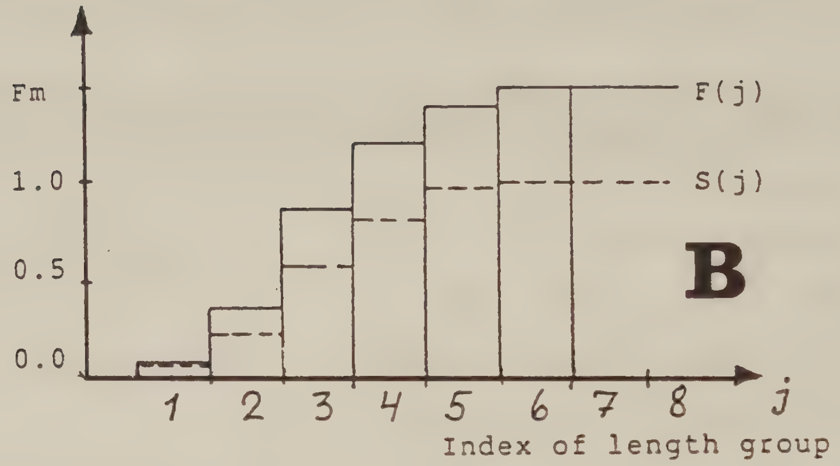


Fig. 6.1.1.3 **Relationship between selection ogive and fishing mortality**
A: Continuous functions
B: The step-functions corresponding to A

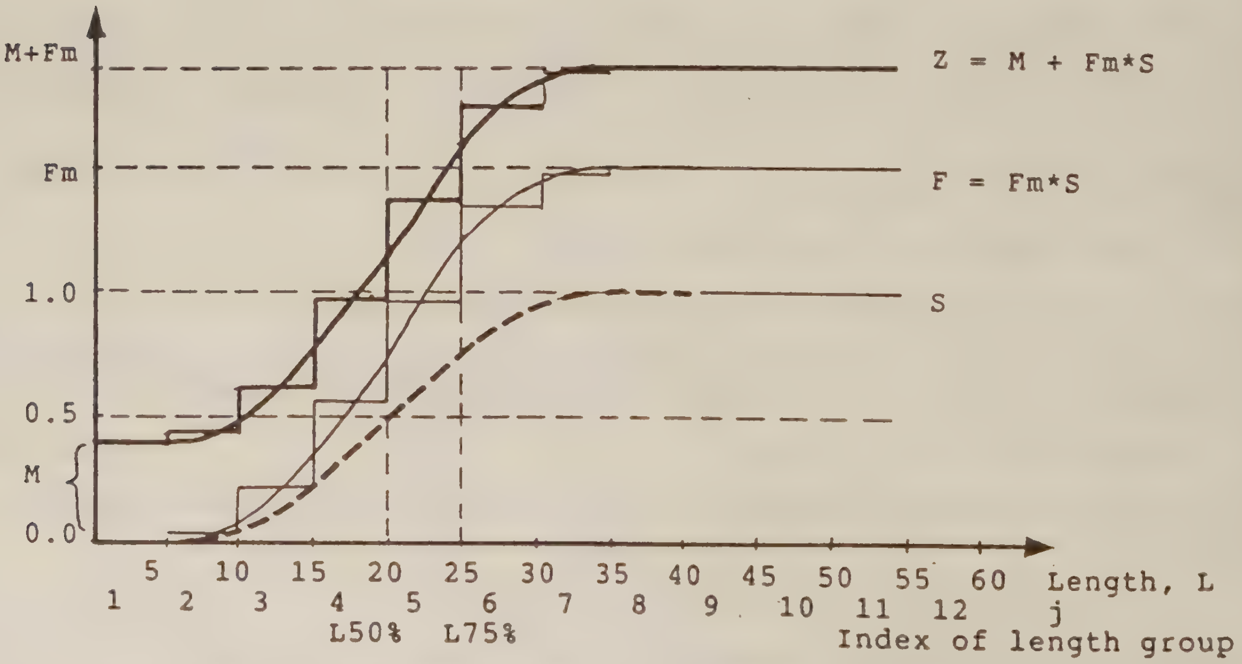


Fig. 6.1.1.4 **The relationship between mortality and the selection ogive (for further explanation, see text)**

real situation. However, knife-edge selection is often used as an approximation to the selection ogive. For lengths below L50% the number selected is underestimated and for lengths above the number is overestimated. These two sources of bias have opposite signs and as the two areas "a" and "b" (see Fig. 6.1.0.1) are the same size they balance out. However, being bigger, the fish of area "a" will weigh more than those of area "b".

So far, selectivity has been considered a function of length. Using the von Bertalanffy growth equation we can express length as a function of age, and express S as a function of age, t:

$$S(t) = \frac{1}{1 + \exp[S1 - S2 * L_\infty * (1-\exp(-K(t-to)))]} \tag{6.1.1.7}$$

The graph of this expression is approximately equal to the graph defined by:

$$S(t) = \frac{1}{1 + \exp(T1 - T2 * t)} \tag{6.1.1.8}$$

where:

$$T1 = t50\% * \ln 3/(t75\% - t50\%) \tag{6.1.1.9}$$

$$T2 = \ln 3/(t75\% - t50\%) = T1/t50\% \tag{6.1.1.10}$$

(Compare Eqs. 6.1.1.4 and 6.1.1.5)

$$t50\% = t_o - (1/K) \ln(1 - L50\%/L_\infty)$$

$$t75\% = t_o - (1/K) \ln(1 - L75\%/L_\infty)$$

Assume that L_∞ = 50 cm, t_o = 0 years and K = 0.3 per year for the fish stock associated with the ogive in Fig. 6.1.1.1. Then:

$$t50\% = 0 - (1/0.3) * \ln(1-15/50) = 1.1889$$

$$t75\% = 0 - (1/0.3) * \ln(1-18/50) = 1.4876$$

$$T1 = 4.3727 \quad \text{and} \quad T2 = 3.6779$$

Fig. 6.1.1.2 shows the exact age transformed selection ogive (curve A, Eq. 6.1.1.7) and the approximated selection ogive (curve B, Eqs. 6.1.1.8-6.1.1.10).

Fishing mortality, F, is clearly related to the selection ogive. When S(L) = 0 fishing mortality must be zero and when S(L) = 1.0 fishing mortality is at its highest level. The obvious relationship between fishing mortality and selection is:

$$F(L) = Fm * S(L) \tag{6.1.1.11}$$

where Fm is the "maximum fishing mortality". Thus, F, as a function of length has the same shape as S, but it has a different level (see Fig. 6.1.1.3A).

In Eq. 6.1.1.11 we consider F a continuous function of length, L. In practice, however, it is often convenient to replace the continuous function by a step function as shown in Fig. 6.1.1.3B. Thus, within each length group we assume F to remain constant.

The continuous selection curve may also be approximated by a step function, in which the value for length group no. j is S((L1+L2)/2), where L1 and L2 are the lower and upper limits of length group no. j. When we use the length group index, j, as argument rather than the length L, we may write a step-function model for total mortality

$$Z(j) = M + Fm * S(j) \qquad (6.1.1.12)$$

where M is the natural mortality coefficient (here assumed to remain constant for all length groups), S(j) is the selection ogive and Fm the maximum fishing mortality. Figure 6.1.1.4 shows F(j), Z(j) and S(j) as functions of length. If Z, M and Fm are known the selection can be estimated by:

$$S(j) = F(j)/Fm \qquad (6.1.1.13)$$

where F(j) = Z(j) - M. (Pope *et al.*, 1975 and Hoydal, Rørvik and Sparre, 1982).

When we work with a step-function rather than the continuous logistic curve the selection is given as an array of S-values and this array can replace the mathematical expression (Eq. 6.1.1.1) or the array may be applied to estimate the parameters of the logistic curve. Actually, an array of S-values is a more versatile way of presenting selection ogives as no assumptions are made about the underlying mathematical expression. (cf. the discussion of age/length keys versus growth equations, Section 3.3.5).

In Section 6.6 it will be shown how selection ogives can be estimated from an array of F-values (the so-called "*fishing pattern*").

(See **Exercise**(s) in Chapter 17).

6.1.2 Estimation of the selection ogive from a catch curve

When using a linearized catch curve to estimate mortality (e.g. Fig. 4.4.5.1) it is usually necessary to discard the left hand side of the curve because the juvenile fish are not fully exploited or not fully recruited. A conceptually simple way to estimate how many fish are missing at each age is to extrapolate on the straight line from which the total mortality coefficent Z is estimated, in order to find the number of juveniles there "ought to be", see Fig. 6.1.2.1. The differences between the "expected" numbers and the actual numbers should give the ogive resulting from the combined effect of recruitment and mesh selection. As shown below, the calculations are easily performed. The problem is that an important and probably unrealistic assumption is made. It was assumed that the total mortality rate, Z = F + M, is the same for all ages. F alone is not constant because it must be smaller in the mesh selection phase. M, on the other hand, is likely to be higher for small fish than for the adults. It is therefore possible that Z remains approximately constant although so far, nobody has shown it to be. Nevertheless, the method has achieved considerable popularity and is therefore related here.

To explain this method (Pauly, 1984a) the example of Table 6.1.2.1 is used. Columns A-E contain the input data and calculations for a length converted catch curve analysis (cf. Section 4.4.5). In this case we calculate the total mortality Z = 1.0 per year from the growth parameters, L_∞ = 50 cm, K = 0.3 per year (See Fig. 6.1.2.1). The result of the regression analysis is:

$$\ln \frac{C}{\Delta t} = 9.208 - 1.0 * t$$

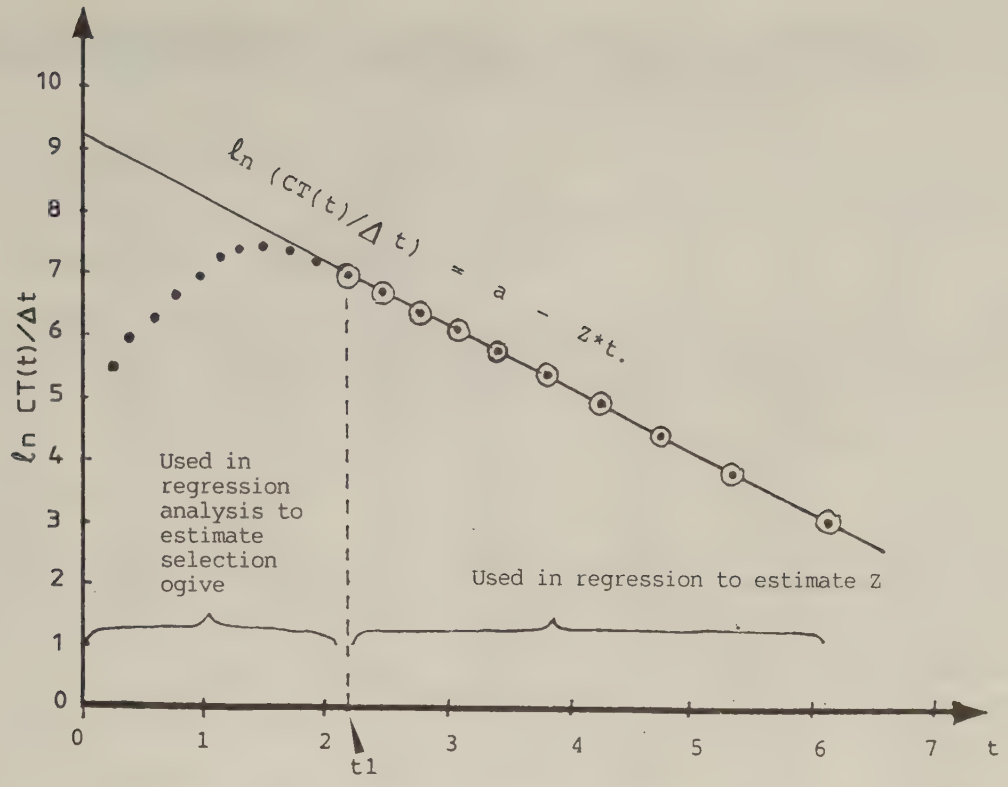


Fig. 6.1.2.1 **Estimation of the selection ogive from a length converted catch curve analysis based on Table 6.1.2.1**

In contrast to the exercise discussed in Section 4.4 we now have a use for the intercept (a = 9.208).

Under the assumption of constant mortality we expect the values of ln(C/$\underline{/\backslash}$t) to be on the regression line, ln(C/$\underline{/\backslash}$t) = a - Z * t. Thus, the hypothetical maximum frequency, CT, is expected to fulfill the equation:

$$\ln (CT/\underline{/\backslash}t) = a - Z * t. \qquad (6.1.2.1)$$

The idea behind this method is that the number in the sea is proportional to the number caught, i.e.

$$\frac{C}{CT} = \frac{\text{The number in the catch}}{\text{Total population number in the sea}}$$

Let t1 be the age corresponding to the first length group used in the catch curve regression (in the case of Table 6.1.2.1, we have t1 = 2.180, see Fig. 6.1.2.1). For ages above t1, CT(t) should be approximately equal to the observed frequencies, since selection and recruitment are supposed to have finished before that age. But for the ages below t1 we expect:

$$\ln (CT(t)/\underline{/\backslash}t) > \ln (C(t)/\underline{/\backslash}t) \qquad \text{(compare 6.1.2.1)}$$

As CT(t) is supposed to be proportional to the population number the ratio

C(t)/CT(t)

is an estimate of the probability that a fish of age t will be on the fishing ground and be retained if it encounters the gear, i.e. C/CT can be used as an estimate of the resultant ogive S.

Table 6.1.2.1 Example to illustrate estimation of the selection ogive from a catch curve (compare Fig. 6.1.2.1). L_∞ = 50 cm, K = 0.3 per year, t_o = 0

A	B	C	D	E	F	G	H
L1-L2	t X	/\t	C(L1,L2)	ln (C//\t) Y'	S obs.	ln (1/S-1) Y	S est.
3- 5	0.278	0.145	37	5.54	0.034	3.35	0.03
5- 7	0.426	0.151	56	5.92	0.057	2.81	0.06
7- 9	0.581	0.159	86	6.29	0.097	2.23	0.10
9-11	0.744	0.167	129	6.65	0.163	1.64	0.16
11-13	0.915	0.176	188	6.97	0.267	1.01	0.27
13-15	1.095	0.186	258	7.23	0.416	0.42	0.42
15-17	1.286	0.196	319	7.39	0.590	-0.37	0.59
17-19	1.487	0.208	352	7.43	0.750	-1.10	0.75
19-21	1.703	0.222	351	7.37	0.870	-1.90	0.87
21-23	1.933	0.238	324	7.22	0.943	-2.80	0.94
23-25	2.180	0.257	283	7.00	(0.976)	-	0.98
25-27	2.447	0.278	239	6.76	-	-	0.99
27-29	2.734	0.303	196	6.47	-	-	1.00
29-31	3.054	0.334	158	6.16	-	-	1.00
31-33	3.406	0.371	123	5.80	-	-	1.00
33-35	3.798	0.417	93	5.41	-	-	1.00
35-37	4.243	0.477	69	4.97	-	-	1.00
37-39	4.757	0.557	48	4.46	-	-	1.00
39-41	5.365	0.669	31	3.84	-	-	1.00
41-43	6.109	0.838	18	3.04	-	-	1.00
43-45	7.068	1.122	10	2.19	-	-	1.00
45-47	8.419	1.702	3	0.57	-	-	1.00

column	content
A	length group in cm
B	t(L1+L2)/2, age corresponding to interval midlength (cf. Eq. 3.3.3.3); independent variable, in both regressions (X)
C	/\t(L1,L2), see Eq. 4.4.5.1
D	C(L1,L2) = catch in numbers per length group
E	ln(C//\t), dependent variable in catch curve regression analysis (Y')
F	S obs = C/[/\t * exp(a - Z * t)], observed selection ogive (a = 9.208, Z = 1.0)
G	ln(1/S-1), dependent variable in regression analysis for selection ogive (Y)
H	S est = 1/[1 + exp(T1-T2 * t)], estimated (theoretical) selection ogive

Note: length range 23 to 43 included in the catch curve regression analysis

CT can be predicted by a modified Eq. 6.1.2.1:

$$CT(t) = \Delta t \;\exp(a - Z * t)$$

Thus, the ogive can be estimated by:

$$S(t) = \frac{C(t)}{CT(t)} = \frac{C(t)}{\Delta t \exp(a - Z * t)} \qquad (6.1.2.2)$$

The fractions retained of the observed selection ogive are presented in column F of Table 6.1.2.1. In order to obtain the theoretical (estimated) selection ogive Eq. 6.1.1.8 is used in linear form.

The expression for S(t) given as Eq. 6.1.1.8 is equivalent to

$$\ln(1/S(t) - 1) = T1 - T2 * t \qquad (6.1.2.3)$$

Eq. 6.1.2.3 enables us to estimate the parameters T1 and T2 by linear regression. Columns B (X) and G (Y) of Table 6.1.2.1 contain the input for this regression. (Columns C, D and E contain the result of the catch curve analysis used to calculate the dependent variable, column G). Column H contains the estimated selection ogive. Only values of S (column F) less than 1.0 can be used in the expression for Y (column G). Carrying out the regression analysis we find:

$$a = T1 = 4.396 \quad \text{and} \quad b = T2 = 3.701$$

which gives:

$$t50\% = T1/T2 = 1.1877 \text{ years}$$
$$t75\% = (\ln 3 + T1)/T2 = 1.4846 \text{ years}$$
$$L50\% = 50 * (1 - \exp(-0.3 * (1.1877-0))) = 15.0 \text{ cm}$$
$$L75\% = 50 * (1 - \exp(-0.3 * (1.4846-0))) = 18.0 \text{ cm}$$

The example of Table 6.1.2.1 is a hypothetical one constructed to give the results of Fig. 6.1.1.1 and 6.1.0.1. Because the data are ideal hypothetical data there is a perfect agreement between the observed ogive (column F of Table 6.1.2.1) and the theoretical selection ogive (column H).

The exercise provides a check of the adequacy of the choice of points used in the regression analysis. The conclusion from Table 6.1.2.1 is that the first length group used in the estimation of Z should have been 27-29 cm, as this group is the first one under full exploitation. However, because the logistic curve never attains the value 1.0 the concept of "full exploitation" is determined by the number of decimals in the table. Taking into account that the logistic curve is an approximation to the real selection curve one cannot expect to get a precise estimate for the first length under full exploitation. If we are somewhere in the "near neighbourhood" of 1.0 the choice of first length group in the catch curve regression is likely to be good enough.

As emphasized in the introduction to this section, the results of the method described should be treated with a certain reservation.

(See **Exercise(s)** in Chapter 17).

6.1.3 Mesh selection experiments

If there is no overlap between the gear selection ogive and the recruitment ogive the gear selection can be estimated by the method described in Section 6.1.2 (cf. Fig. 6.1.0.3), bearing in mind the restrictive assumptions underlying the method.

A more direct method is to set up a gear selection experiment. The most commonly used experiment is the "covered codend method" (described by among others Pope et al., 1975 and Jones, 1976). The idea behind the experiment is illustrated in Fig. 6.1.3.1. The relatively large meshed codend of the trawl is covered with a small meshed net, which retains (almost) all fish. By comparing the length compositions of the fish retained in the codend and in the cover, the probability of escapement through the large-meshed net can be estimated.

Table 6.1.3.1 and Fig. 6.1.3.2 show an example of such a covered codend experiment (from Jones, 1976). The experiment deals with the threadfin bream, Nemipterus japonicus, caught with a trawl with a codend mesh size of 4 cm. Only the length groups from 10 cm to 17 cm are used in the analysis, because the observed retention percentages of 0% or 100% cannot be used. The model to estimate the selection ogive is defined by Eq. 6.1.1.1 rearranged into

$$\ln(1/S(L) - 1) = S1 - S2 * L \qquad (6.1.3.1)$$

This is the length-based counterpart to the age-based model applied in Section 6.1.2 (Eq. 6.1.2.3).

Mathematically Eq. 6.1.3.1 and Eq. 6.1.2.3 are identical. They differ only as far as the unit of the independent variable, L, is concerned. L is the length interval midpoint and S(L) is the fraction of length group L fish retained in the codend, i.e.:

$$S = \frac{\text{number of fish in the codend}}{\text{number of fish in the codend and in the cover}}$$

It is seen that if S = 0 or if S = 1 the expression in Eq. 6.1.3.1 is not defined.

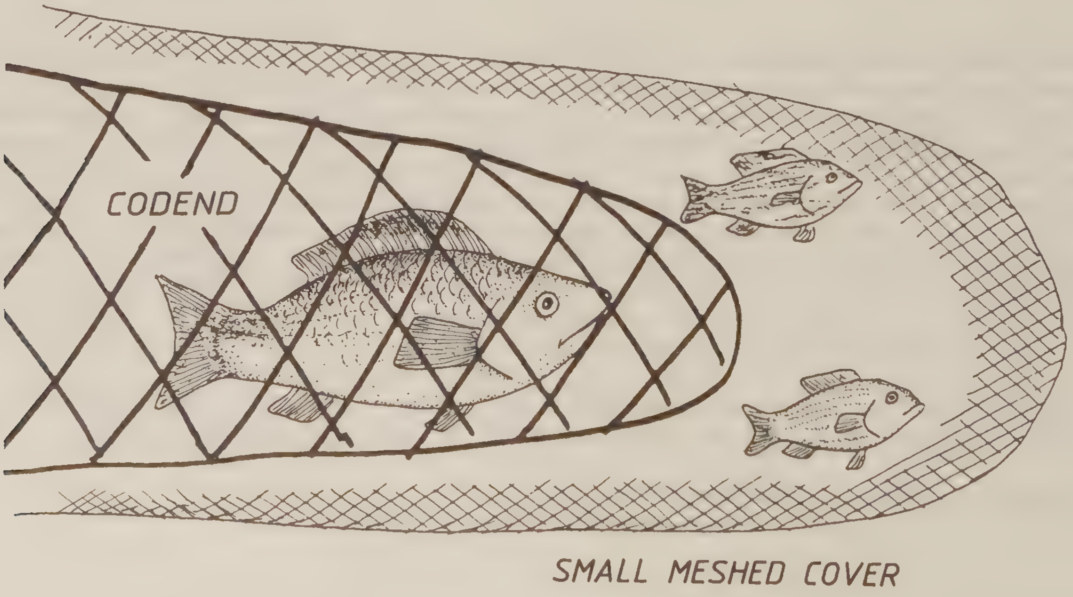


Fig. 6.1.3.1 Covered codend

Table 6.1.3.1 **Estimation of gear selection ogive for Nemipterus japonicus from a covered codend experiment (from Jones, 1976, compare Fig. 6.1.3.2)**

length interval L1-L2	number in codend	number in cover	total number	fraction retained S	Y ln(1/S-1)	X midlength (L1+L2)/2
9-10	0	1	1	0	-	-
10-11	1	6	7	0.14	1.82	10.5
11-12	2	7	9	0.22	1.27	11.5
12-13	2	4	6	0.33	0.71	12.5
13-14	7	5	12	0.58	-0.32	13.5
14-15	30	13	43	0.70	-0.85	14.5
15-16	61	8	69	0.88	-1.99	15.5
16-17	27	3	30	0.90	-2.20	16.5
17-18	7	0	7	1.00	-	-
18-19	4	1	5	0.80	-	-

intercept = S1 = 9.4875 -slope = S2 = 0.7193

L50% = S1/S2 = 13.2 cm L75% = (ln 3 + S1)/S2 = 14.7 cm

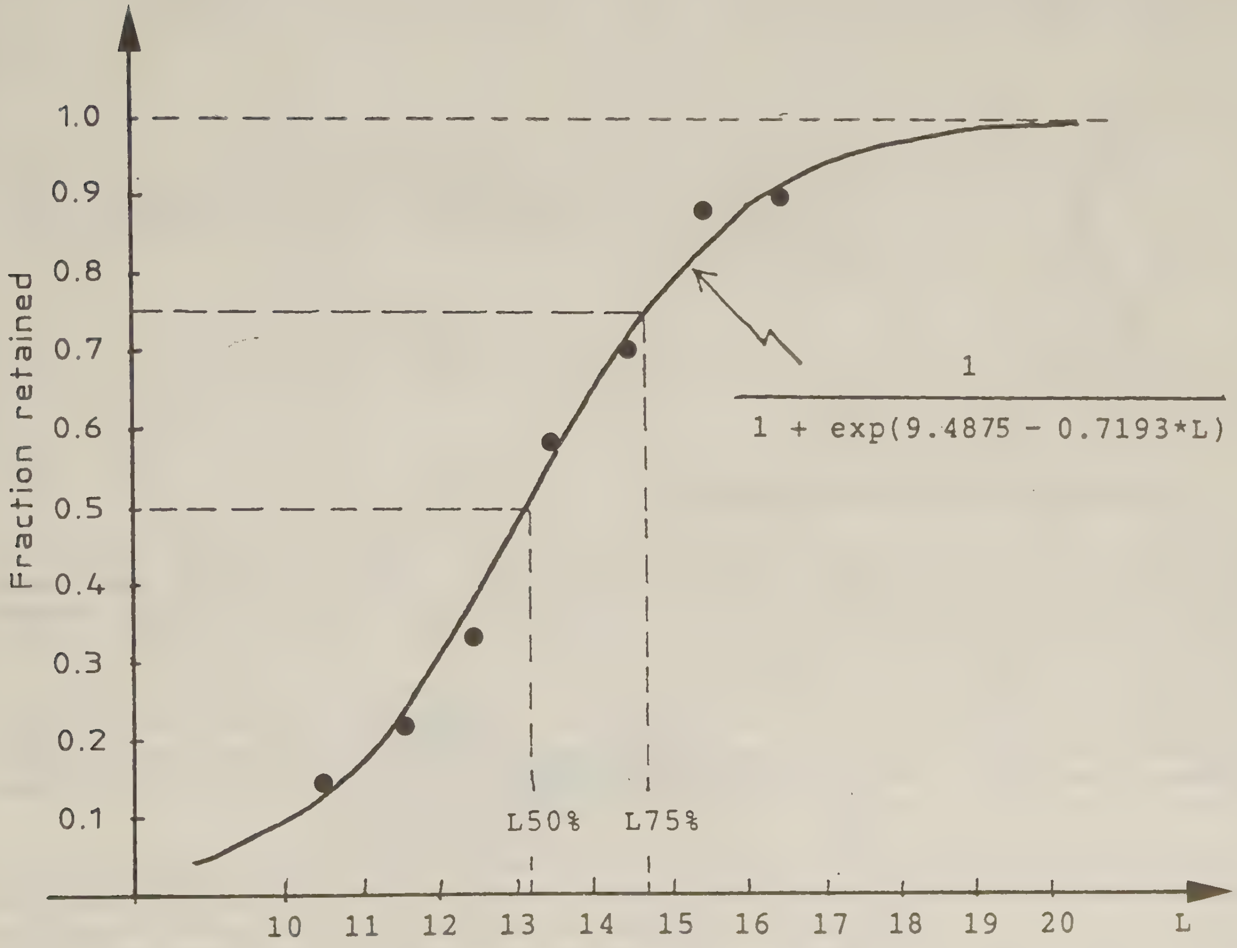


Fig. 6.1.3.2 **Gear selection ogive for Nemipterus japonicus caught by a 40 mm trawl (from Jones, 1976)**

The estimation procedure is illustrated in Table 6.1.3.1. The regression analysis (Eq. 6.1.3.1) gives the results:

Intercept: $S1 = 9.4875$, Slope: $- S2 = - 0.7193$ which gives:

$$L50\% = S1/S2 = 13.2 \text{ cm}$$
$$L75\% = (\ln 3+S1)/S2 = 14.7 \text{ cm}$$

(cf. Eqs. 6.1.1.2 and 6.1.1.3). Fig. 6.1.3.2 shows the observed retained fractions together with the estimated gear selection ogive.

As the probability that a fish will escape through the meshes depends on the body depth of the fish compared to the mesh size, it is natural to assume proportionality between d50% (the body depth at which 50% of the fish are retained) and the mesh size:

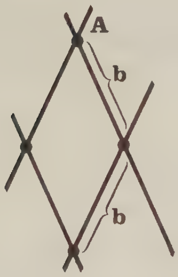
$$d50\% = A * (\text{mesh size}) \tag{6.1.3.2}$$

where A is a constant. As body depth is approximately proportional to body length (cf. Section 2.4) Eq. 6.1.3.2 implies that a similar expression holds for the length of the fish:

$$L50\% = S.F. * (\text{mesh size}) \tag{6.1.3.3}$$

The factor of proportionality, S.F. is called the "selection factor". In the case of Table 6.1.3.1 we found L50% = 13.2 cm for a mesh size of 4 cm. Thus, the selection factor is:

$$S.F. = 13.2/4 = 3.3$$



The "mesh size" is usually defined as the length of the "stretched" whole mesh. The mesh size of the netting shown here is 2 * b, where b is the length between two knots.

For a detailed discussion of definitions of mesh size and mesh measuring techniques, see FAO (1978b).

6.2 Estimation of gill net selection

Basically, the gill nets are rectangular nets where the upper edge, the head rope has floats while the foot rope has sinkers. Often gill nets (drifting and set nets) are in the form of gangs of nets with different mesh sizes. For further descriptions of gill nets, see FAO (1978b), Nédélec (1982) or Karlsen and Bjarnason (1986).

The selection properties of gill nets are reviewed in Hamley (1975). Discussions on gill net selectivity can be found in e.g. Baranov (1948); McCombie and Fry (1960), Gulland and Harding (1961), Regier and Robson (1966), Hamley and Regier (1973) and Jensen (1986).

They are "passive gears", i.e. the fish has to swim into the net to get caught whereas trawls are "active gears". Theoretically, this implies that fish which move fast, have a larger probability of encounter with the passive gear than slow moving fish. Large fish move faster than small fish of the same species. The swimming speed can be approximated by a constant times a power function of length:

$A * L^B$, where A and B are constants (Yates, 1983).

Rudstam, Magnuson and Tonn (1984) included swimming speed (with B = 0.8 for the cisco, Coregonus artedii, from Wisconsin, USA) into a model for gill net selection. They considered the selection as the product of two probabilities:

(selection) = (probability of encounter) * (probability of being caught given encounter)

We shall, however, in the following deal only with the last factor, the probability of being caught given encounter.

For simple gill nets the selection curve has (unlike trawl selection) a right hand side descending slope. Small fish can pass through the meshes as was the case for trawl but large fish may also avoid being caught in a gill net, because their heads are so large that they cannot be "gilled". This is the simple theory behind gill net selection. The picture becomes somewhat more complicated when other ways by which the fish can be stuck in a gill net are also considered. Baranov (1914) recognized three ways (see Fig. 6.2.1)

a. Gilled

b. Wedged, i.e. held tightly by a mesh around the body

c. Tangled, i.e. held in the net by teeth, maxillaries, fins or other projections, without necessarily penetrating the mesh.

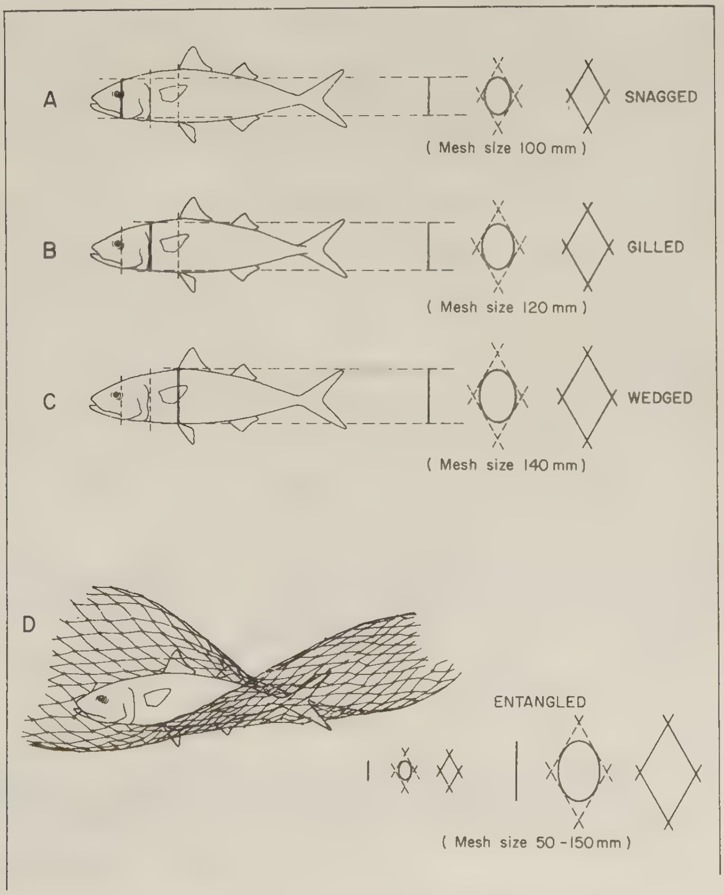


Fig. 6.2.1 Gill net selection

Table 6.2.1 **Estimation of gill net selection ogives for Tilapia esculenta, Lake Victoria (from Garrod, 1961)**

length interval midpoint L X	number caught mesh 8.1 cm Ca	number caught mesh 9.1 cm Cb	$\ln \frac{Cb}{Ca}$ Y
18.5	7	0	not
19.5	90	1	used
20.5	199	9	-3.096
21.5	182	53	-1.234
22.5	119	290	0.891
23.5	29	357	2.510
24.5	17	225	-
25.5	3	82	not
26.5	0	19	used
27.5	0	10	-

intercept = - 41.907, slope = 1.894

$$Lma = -2 * \frac{(-41.907) * 8.1}{1.894 * (8.1+9.1)} = 20.8 \text{ cm}$$

$$Lmb = 20.8 * 9.1/8.1 = 23.4 \text{ cm}$$

$$s = \sqrt{\frac{2 * (-41.907) * (8.1+9.1)}{1.894^2 * (8.1 + 9.1)}} = 1.17$$

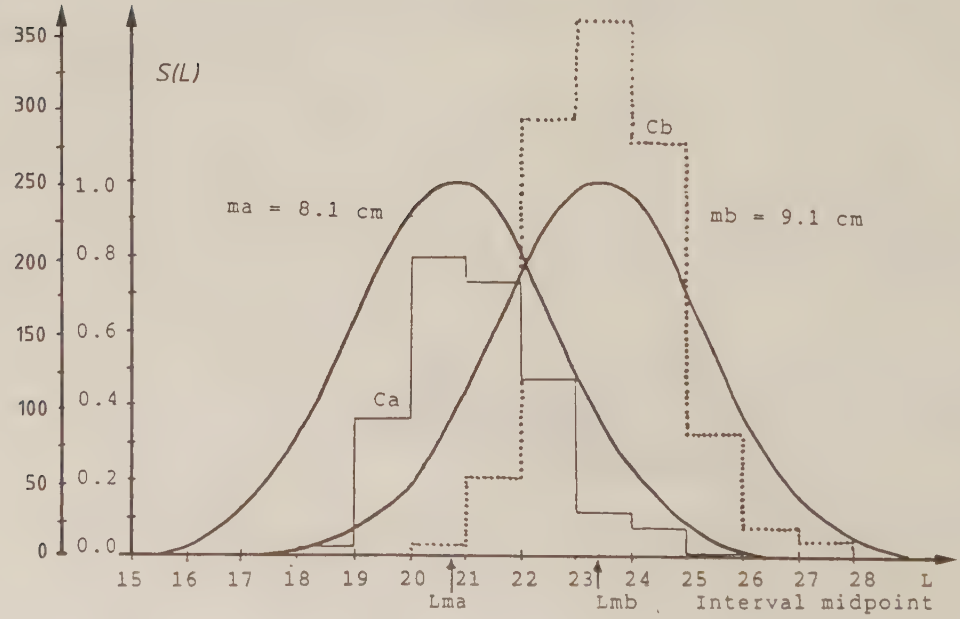


Fig. 6.2.2 **Selection curves for Tilapia esculenta from Lake Victoria (from Garrod, 1961)**

For the two first types of getting stuck in a gill net it has been suggested (Holt, 1963) that the selection curves are bell shaped and can be described by the normal distribution (cf. Section 2.2). Thus for gilling or wedging we use the model:

$$S(L) = \exp\left[-\frac{(L - Lm)^2}{2 * s^2}\right] \quad (6.2.1)$$

where Lm is the "optimum length for being caught" and s is the standard deviation of the normal distribution. The factor "$n * dL/(s\sqrt{2\pi})$" in the expression for the normal distribution (Eq. 2.2.1) is not used here. By leaving out this factor, S becomes a fraction, i.e. $0 < S <= 1.0$

Holt (1963) suggested an experiment to estimate Lm and s by using two gill nets with different mesh sizes, ma and mb. The two mesh sizes must be such that their selection ogives overlap. The assumptions behind the method are:

1. The optimum length Lm (the top of the bell-shaped selection curve) is proportional to the mesh size (i.e. Lm = S.F. * m, where S.F. is the selection factor, cf. Section 6.1.3)
2. The two selection curves have the same standard deviation
3. The two gears have the same fishing power. This includes that when set, they must have the same area.

The two nets are used to fish in the same area at the same time and the observations are numbers caught by length group.

Table 6.2.1 columns A, B and C show an example of such an experiment on Tilapia esculenta from Lake Victoria. Column B contains the number caught, Ca, by length group for the smaller meshed gear and column C contains the corresponding numbers, Cb, for the larger meshed net. The parameters to be estimated are:

Lma: Optimum length for the smaller meshed net

Lmb: Optimum length for the larger meshed net

s: The common standard deviation

Input data for the analysis are numbers caught by length group for each gear, Ca and Cb, and the two mesh sizes ma and mb. The mathematical derivations are lengthy and are by-passed here. The first step in the analysis is to calculate the log ratios

$$Y = \ln (Cb/Ca)$$

for each length group (column D of Table 6.2.1). Only the lengths where the frequencies overlap can be used. The second step is to do a regression analysis of the log ratios (Y = ln(Cb/Ca), column D) against the interval midpoint (X = L, column A):

$$\ln (Cb/Ca) = a + b * L \quad (6.2.2)$$

The third step gives the results (see example in Table 6.2.1). The optimum length for the smaller meshed gear is determined by:

$$Lma = - 2 * \frac{a * ma}{b * (ma + mb)} \quad (6.2.3)$$

and the optimum length for the larger meshed gear by:

$$Lmb = -2 * \frac{a * mb}{b * (ma + mb)} = Lma * mb/ma \qquad (6.2.4)$$

The common standard deviation is:

$$s = \sqrt{\frac{2 * a * (mb - ma)}{b^2 * (ma + mb)}} \qquad (6.2.5)$$

The selection factor S.F. is estimated from:

$$S.F. = \frac{-2 * a}{b * (ma + mb)} \qquad (6.2.6)$$

Fig. 6.2.2 shows the two estimated selection curves for Tilapia esculenta, calculated by means of Eq. 6.2.1. Introducing the selection factor, Eqs. 6.2.3 and 6.2.4 can be written:

$$Lma = S.F. * ma \quad \text{and} \quad Lmb = S.F. * mb$$

When more than two mesh sizes are fished the estimates of an overall selection factor and a common standard deviation can be obtained from the results of each pair of successive mesh sizes, under the same assumptions as for the case with two mesh sizes. Assume that there are n mesh sizes. This gives n-1 estimates of the intercept, a, and slope, b, in Eq. 6.2.2. Thus we have the results:

$$[a(1),b(1)],[a(2),b(2)],....,[a(n-1),b(n-1)]$$

corresponding to the mesh sizes:

$$[m(1),m(2)],[m,2),m(3)],....,[m(n-1),m(n)]$$

The next step is to estimate the selection factor by making a regression analysis through the origin (cf. Section 2.4) with y(i) = -2 * a(i)/b(i) as dependent variable and m(i)+m(i+1) as independent variable. The slope then becomes the selection factor as (cf. Eq. 6.2.6):

$$-2a(i)/b(i) = S.F. * [m(i)+m(i+1)], \; i = 1.2,...,n-1 \qquad (6.2.7)$$

i.e. we get the overall selection factor from Eq. 2.4.13:

$$S.F. = -2 \sum_{i=1}^{n-1} [a(i)/b(i)] * [m(i)+m(i+1)] \Big/ \sum_{i=1}^{n-1} [m(i)+m(i+1)]^2$$

The common standard deviation is estimated as the mean value of the individual estimates for each consecutive pair of mesh sizes, i.e.:

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n-1} \frac{2 * a(i) * [m(i+1) - m(i)]}{b(i)^2 * [m(i)+m(i+1)]}}$$

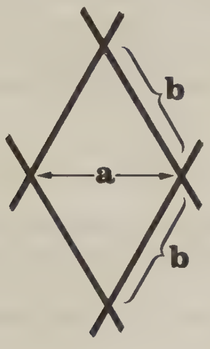
Optimum length for mesh size i is finally obtained by:

$$Lmi = S.F. * mi \qquad (6.2.8)$$

When gangs of gill nets with different mesh sizes are used the combined selection curve for all mesh sizes may resemble the trawl selection ogive (the logistic curve, Eq. 6.1.1) because the fishermen will have ensured the capture of large fish by including large-meshed nets whereas meshes suitable for the smallest fish are not used.

So far we have considered selection when the fish are caught by gilling or wedging only, which we described by a bell-shaped selection curve for a particular mesh size. This type of selection is observed to depend on mesh size and a number of other factors such as net construction, visibility of net, stretchability of net, net material, shape and behaviour of the fish (Hamley, 1975).

Tangling, however, may be less dependent on mesh size and more dependent on other factors. The selection pattern combined for gilling, wedging and tangling may deviate from the bell shaped curve. As only the fish too large to get gilled or wedged are likely to get tangled, the combined selection curve would become a unimodal curve skewed to the right hand side, in the extreme a logistic curve (i.e. selection curve of a trawl). The probability of a fish being tangled is believed to depend on the so-called "hanging ratio", (or "hanging coefficient") which is defined (FAO, 1978b) as:



$$\frac{\text{Length of the head rope}}{\text{(number of meshes) * (mesh size)}}$$

or (see figure):

$$\text{hanging ratio} = \frac{a}{2 * b} \qquad (6.2.9)$$

Thus, for a square mesh ($b = a/\sqrt{2}$) we have the hanging ratio $\sqrt{2}/2 = 0.707$ which represents the maximum opening. Hanging ratios are usually in the range from 0.2 to 0.7. The smaller the hanging ratio the larger the probability of entangling. This is demonstrated by Riedel (1963), who reported catches of Tilapia mossambica with 10 cm mesh gill-nets of three different hanging ratios:

Hanging Ratio	Average number caught per day	Percent tangled	Size range of 95 % of catch
0.707	9.3	0	18-23 cm
0.36	29.5	24	13-23 cm
0.24	81.0	80	8-22 cm

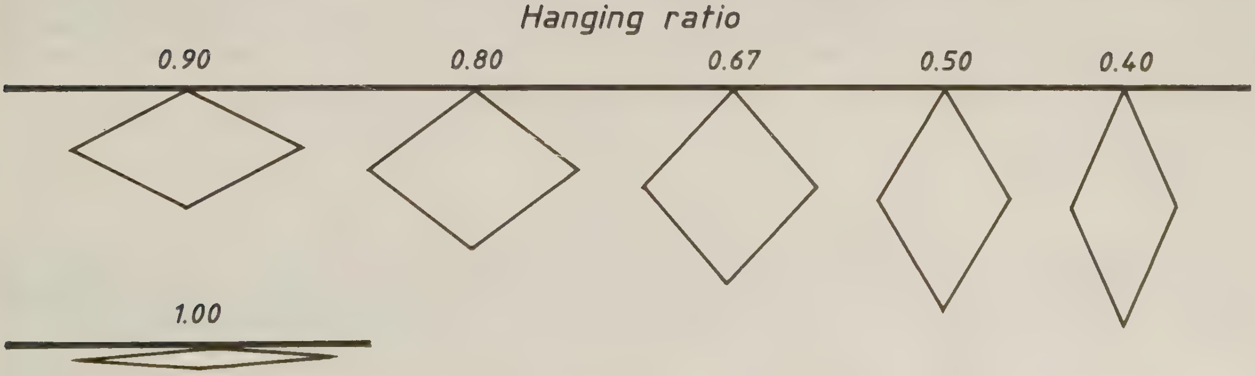


Fig. 6.2.3 Mesh shapes with various hanging rations for gill nets

If entangling is an important factor, the normal distribution method described above is not suitable. One way of estimating the selection curve would be to compare the catches of the gill net to a non-selective gear, e.g. to trawl catches. The catches from the non-selective gear would then play the same role as the total catch in cover and codend of a trawl when operated as described in Section 6.1.3. The same procedure as used in Table 6.1.3.1 can be applied. In this case we need a non-symmetrical selection curve of the type shown in Fig. 6.2.4. A mathematical expression for this type of curve can be obtained by multiplying two logistic curves (Hoydal, Rørvik and Sparre, 1982). The ascending part of the curve is given by the usual logistic curve (cf. Eq. 6.1.1.1). This part of the curve reflects the probability of being gilled or wedged. We call it "SL" where "L" stands for "left hand side of the selection curve":

$$SL(L) = 1/(1 + \exp(S1-S2 * L)) \qquad (6.2.10)$$

This type of selection is the dominating one until length A (see Fig. 6.2.4). After length B the selection is the combined effect of gilling, wedging and entangling. This part of the curve is modeled by a "reversed logistic curve". We call it "SR" where "R" stands for "right hand side of the selection curve":

$$SR(L) = 1/(1 + \exp(D1-D2 * L) \qquad (6.2.11)$$

The parameters in the SR-function, D1 and D2, are negative numbers, whereas the parameters in the SL-function, S1 and S2, are positive numbers. The lengths corresponding to the 50% and 75% de-selection, D50% and D75%, are related to D1 and D2 by the same mathematical expressions as used for S1, S2, L50% and L75% (cf. Eqs. 6.1.1.2-6.1.1.5):

$$D50\% = D1/D2 \quad \text{and} \quad D75\% = (\ln 3 + D1)/D2$$

$$D1 = D50\% * \ln 3/(D75\% - D50\%) \quad \text{and} \quad D2 = D1/D50\%$$

By multiplying the ascending curve, SL, and the descending curve, SR, we obtain the desired type of curve, S. When SL is ascending SR is "neutral" (i.e. approximately equal to 1.0), and when SR is descending SL is "neutral":

$$S(L) = SL(L) * SR(L) = \frac{1}{1+\exp(S1-S2 * L)} * \frac{1}{1+\exp(D1-D2 * L)} \qquad (6.2.12)$$

The expression for S(L) is supposed to take the maximum value, 1.0, for at least one L-value. Therefore, the expression Eq. 6.2.12 (as well as the normal distribution in Eq. 6.2.1) should be normalized so that the maximum value equals 1.0.

In practice this is obtained in the following way. Let S(i) be the point on the selection curve representing length class no. i (estimated by Eq. 6.2.1 or Eq. 6.2.12) and let MAX { S(j) } designate the maximum value of S(j) among all length classes. We "normalize" by replacing the value of S(i) by the value:

$$\frac{S(i)}{\underset{j}{\text{MAX}}\ \{ S(j) \}}$$

The parameters can be estimated as the parameters of the trawl selection ogive (cf. Section 6.1.3). To estimate S1 and S2 we use only the length classes below A (see Fig. 6.2.4) and perform the regression analysis:

$$\ln(1/S-1) = S1 - S2 * L \qquad (6.2.13)$$

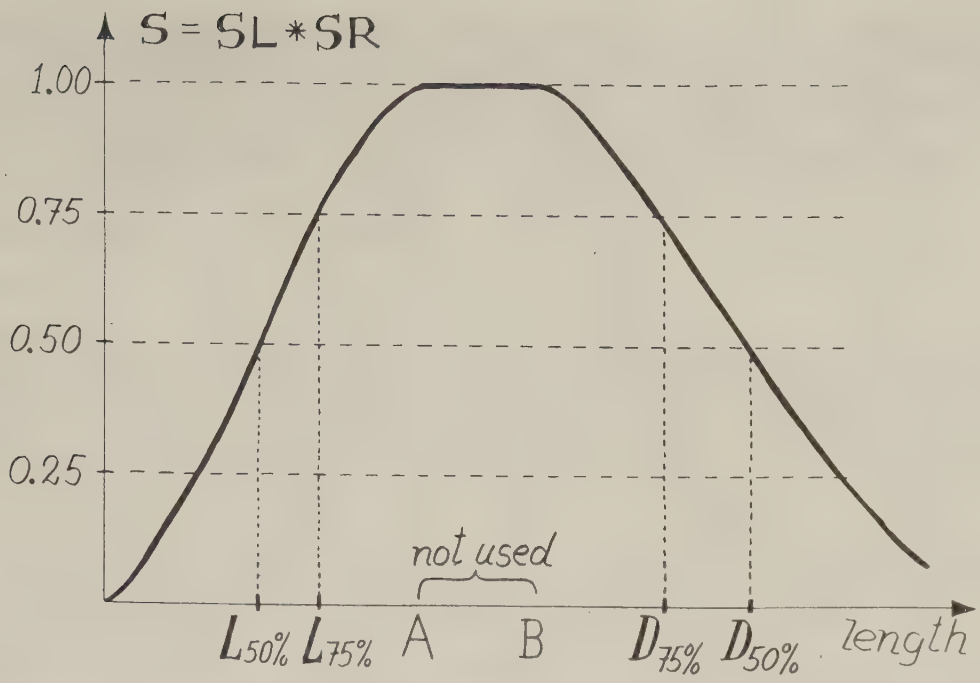


Fig. 6.2.4 Asymmetric selection curve

where the independent variable, ln(1/S-1), is derived from the comparison to the non-selective gear:

$$S(i) = \frac{Cg(i)/Cn(i)}{\underset{j}{MAX}\ \{ Cg(j)/Cn(j) \}} \tag{6.2.14}$$

$$\frac{Cg(i)}{Cn(i)} = \frac{\text{no. of length class i fish caught in gill net}}{\text{no. of length classes i fish caught in non-selective gear}}$$

where Cg(i) is the number in length group i of the sample from the gill net and Cn(i) the corresponding number in the non-selective gear (e.g., small-meshed trawl).

The denominator in Eq. 6.2.14, "MAX { Cg(j)/Cn(j) }", is the maximum value of the ratio Cg/Cn among all length groups with non-zero values of Cg and Cn. Thus, S as defined by Eq. 6.2.14 takes values between 0 and 1 (including 1).

Eq. 6.2.14 is based on the assumption that the numbers Cn are caught by a gear which is non-selective for all length classes caught by the gill-net. If this is not the case Eq. 6.2.14 should be replaced by:

$$S(i) = \frac{[Cg(i)/Cn(i)] * Sn(i)}{\underset{j}{MAX}\ \{ (Cg(j)/Cn(j)) * Sn(j) \}} \tag{6.2.15}$$

where Sn(i) is the selection curve for the "other gear".

The descending (right hand side) of the selection curve is estimated using the same calculation procedure and the corresponding observations as those of the left hand side. Thus we use the data for the length classes above point B (see Fig. 6.2.4) and perform the linear regression analysis:

$$\ln(1/S-1) = D1 - D2 * L \tag{6.2.16}$$

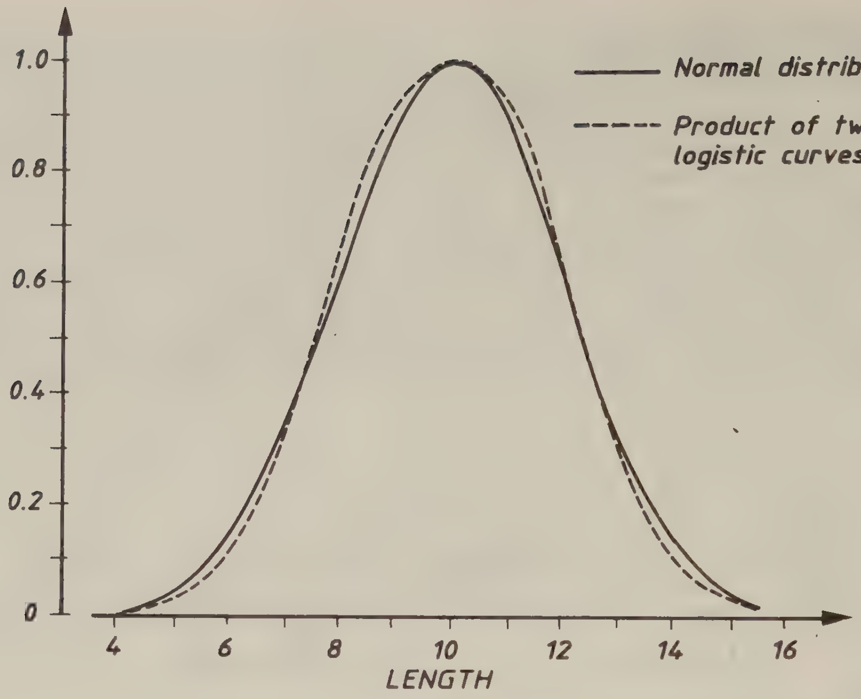


Fig. 6.2.5 A product of two logistic curves which is nearly identical to a normal distribution. The parameters are:
Normal distribution: Lm = 10 and s = 2
Product of two logistic curves:
L50% = 7.645 L75% = 8.483
D50% = 2 * Lm-L50% D75% = 2 * Lm-L75%

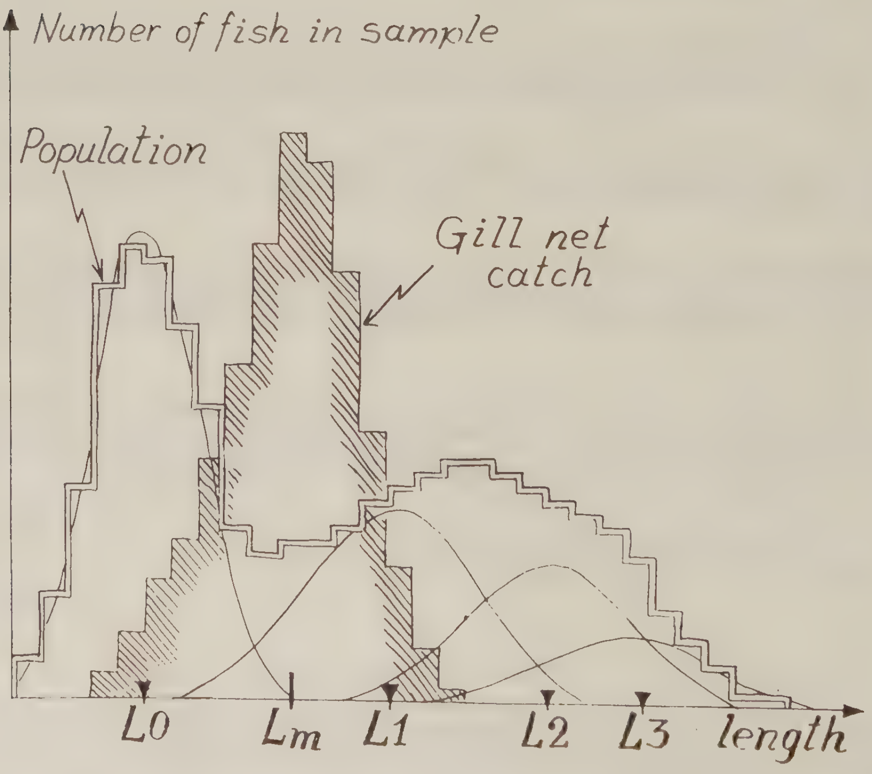


Fig. 6.2.6 Hypothetical example to illustrate bias problems when using gill net data for estimation of growth parameters and mortality rates

The above given method based on the product of two logistic curves is a generalization which includes the trawl selection ogive (Eq. 6.1.1.1) and the symmetric curve (Eq. 6.2.1) as special cases. Assigning the values $D1 = -\infty$ and $D2 = 0$ makes the factor $1/(1+\exp(D1-D2 * L))$ in Eq. 6.2.12 take the value 1 for all values of L and Eq. 6.2.12 equals Eq. 6.2.1). If the curve is symmetrical we estimate parameters so that:

$$L50\% + D50\% = L75\% + D75\%$$

but do not have to make assumptions beforehand.

The symmetrical curve created by the product of two logistic curves may not be the same as the normal distribution curve. The product of two logistic curves may take the maximum value (1.0) for a range of values (cf. Fig. 6.2.4). It may also, however, come very close to the normal distribution as shown by an example in Fig. 6.2.5.

Thus, if data for a non-selective gear (or a gear with known selection curve) are available, there is really no need to use the traditional model, the normal distribution, as the same curve can be obtained as a special case of the logistic product function. Moreover, the latter is versatile and easy to handle from a computation point of view. Using the product of two logistic curves we are not forced to make the questionable assumption that the selection curves are normally distributed with a common standard deviation.

Finally, it should be mentioned that data collected from gill net catches are difficult to use for estimation of growth parameters or mortality rates, as illustrated in Fig. 6.2.6. In this (hypothetical) example we consider a length frequency sample representing the population (double line in Fig. 6.2.6) i.e. a sample from a non-selective gear and a sample representing the gill net catch (shaded line in Fig. 6.2.6). The population contains four components, whereas the gill net sample appears to contain only one component with mean value Lm, which does not coincide with any of the four cohort mean lengths (L0,L1,L2 and L3). Thus, the gill net sample does not give any information which can be used for the estimation of length by age.

Samples collected during the year would all (more or less) give the same picture. Summing a time series over a year would therefore give a curve not very different from the curve for the gill net catch in Fig. 6.2.6. Using the descending slope of this curve to estimate total mortality from the length based catch curve analysis would lead to an overestimate. Before any gill net data are used for estimation of growth parameters or mortality rates, they should be critically examined. The result of such an examination may be either that the data cannot be used or that they can be adjusted for gear selection as described in Section 6.5, and subsequently used for estimation of parameters.

6.3 Discussion of selection by other gears

The previous two sections dealt with the selectivity of trawls (active gears) and gill nets (passive gears). This is because the scientific literature on selectivity has been mainly concerned with these gears and because it is relatively easy to conduct experiments for estimation of their selection curves. Other types of gear are also more or less selective and the pattern of selection can usually be changed by suitable adjustments of the gear. The gear selection model based on the logistic curve (Eq. 6.1.1.1) or the model based on the product of two logistic curves (Eq. 6.2.12) are believed versatile enough to describe the selection curve of any gear.

Below follows a brief discussion of the selection properties of two passive gear types: hooks and traps, and of an active gear type: seine nets.

Much more appears to be known about trawl and gill net selection than about hook fisheries. For hook-and-line fisheries some authors report that the selection curve is bell-shaped, of the gill net type, dependent on hook size, whereas others find a trawl type selection. The idea behind the use of the bell-shaped selection curve is that small fish cannot take a large hook in their mouth and large fish are not held securely by small hooks.

A discussion of hook selection can be found in Ralston (1982), who observed the trawl type of selection for the Hawaiian deep-sea hand line fishery. From an experiment where four hook sizes were used for catching snappers and groupers he found that small hooks were nearly as efficient as large hooks at catching large fish. For the ascending left hand selection ogive (the small fish) he found a sigmoid curve. Pope et al. (1975) suggested to use the gill net type of selection for hook fisheries, although they quote a number of works which show the trawl type of selection.

Pope et al. (1975) suggested the trawl type of selection for traps arguing that traps behave like codends in retaining fish. Munro (1974, 1983) discusses selectivity and other aspects of the operation of portable Antillean fish traps. These 3 m long traps are made of chicken cage wire with two funnels whose bottoms are the entrance openings. They are used for catching coral reef fish. Munro (1974) developed a model for the catchability of traps as a function of the soak time (the time the trap is left on the fishing ground).

Trap selectivity is complicated because it relies on the fish to move actively into the trap. For small traps (as the Antillean trap mentioned) only one specimen of a territorial fish species can be expected to be caught, but the trapped fish may have been replaced by another before the trap is pulled. As the other extreme we have the hunters, like the jacks (Carangidae) in a coral reef fishery which cover a lot of ground in a short time and therefore become over-represented in trap catches. If a large predator is caught in the trap it may keep potential prey fish from entering the trap. If prey species are already in the trap they may act as live bait and attract the large predator, which in turn may eat the prey before the trap is pulled.

Trap catches depend on the duration of the soak (Munro, 1974). There is always a chance that a trapped fish finds the entrance opening and leaves through it. There are considerable differences between species. Some species leave the trap with great ease (Munro, 1983). Thus, trap selectivity may not only be a function of the mesh size used in the trap. For example, the size of the entrance opening and the soak time may be of importance. The species composition on the ground where the trap is placed also may influence the selectivity. For the escapement through the meshes, however, it appears reasonable to assume the trawl selection type of ogive. When considering the average of a large number of trap catches some of the above mentioned complications may disappear, i.e. they may turn out to be "random noise" around the selection curve.

In principle a seine net should work like a trawl as far as selection is concerned. However, it is more difficult to deal with the seines because this type of gear is usually used to catch schooling species, such as sardines, mackerels and tunas. These species have a tendency to form schools consisting of fish of the same size. Thus, we should consider a school to be a sampling unit (and not an individual fish).

6.4 Estimation of selection curves from cohort analysis

The various kinds of cohort analyses (Chapter 5) produce an array of estimates of the coefficient F of fishing mortality, by age group or by length groups. These F values supply data for a gear selectivity/recruitment curve which is obtained by:

$$S(i) = \frac{F(i)}{\underset{j}{\mathrm{MAX}}\ \{ F(j) \}} \qquad (6.4.1)$$

where F(i) is the fishing mortality for age or size group i and MAX {F(j)} is the maximum value of the fishing mortality among all age or size groups (cf. Eq. 6.1.13). Eq. 6.4.1 applies to any gear or any combination of gears combined with any recruitment curve (Hoydal, Rørvik and Sparre, 1980 and 1982). The method makes no assumption on the type of gear or on how the fish are caught. Thus, the selection curve of a gear can be estimated from catch data only.

The approach has a number of advantages compared to selection curves derived from measurement of gear characteristics, e.g. mesh size. Let us for example consider the trawl. If we assume (as is often done) that the gear selection curve is determined by the mesh size in the codend only, then two fishing boats using gears with the same mesh size should have the same gear selection ogive. However, this is likely to be the case only if the two boats operate the gear exactly the same way. For example if one boat takes hauls of duration 5 hours and the other boat only uses one hour for a haul the selective properties may be different because of clogging of the net by the catch. Also the towing speed may influence the selectivity. Higher speed may make the meshes more elongated and cause a lower selection factor.

Thus, Eq. 6.4.1 gives the "effective mesh size" (Hoydal, Rørvik and Sparre, 1982), i.e. the observed selection/recruitment parameters. The concept of effective mesh size also applies to gear without meshes, such as hook and lines.

6.5 Using a selection curve to adjust catch samples

When analysing a length frequency sample, (e.g. when doing the Bhattacharya analysis, Section 3.4.1) selection may create biased results. As an example we look at the first part of Table 3.2.1.1. The sample is shown in column B of Table 6.5.1. Actually column B is the first component estimated by the Bhattacharya analysis as can be seen from column H of Table 3.4.1.1, which was based on the same data. These hypothetical data were supposed to represent a random sample of the population. Thus, in the case of Table 3.2.2.1 we assumed a non-selective gear. If the sample had been taken with a selective gear it would have been different.

Suppose that a gear with a trawl type selection with L50% = 15 cm and L75% = 18 cm had been used. In that case we would have observed the frequencies shown in column C of Table 6.5.1 and not those of column B. Figures in column C are hypothetical ones and are calculated as the product of column B and column D (the fraction retained). If selection occurs, column C would be the observations. The frequencies of column C yield a biased estimate of the mean length and the standard deviation as can be seen from the last two rows of Table 6.5.1. However, if the selection ogive is known it is possible to estimate the unbiased sample, i.e. to estimate column B. This is done by dividing the observed frequencies (column C) by the fractions retained. This raising procedure gives column E. As could be expected there are problems with small frequencies (lengths 12-14 cm). The method cannot be used to raise a zero frequency and is not dependable for small frequencies.

Table 6.5.1 Example to illustrate the estimation of a random sample from a sample biased by selection (compare Fig. 6.5.1)

A	B	C	D	E
"true" length group	observed unbiased sample	selection biased sample	estimated ogive S	estimated unbiased sample C/D
12-13	1	0	0.30	0
13-14	4	1	0.37	3
14-15	11	5	0.45	11
15-16	24	13	0.55	24
16-17	38	24	0.63	38
17-18	42	30	0.71	42
18-19	33	26	0.78	33
19-20	20	17	0.84	20
20-21	7	6	0.88	7
21-22	2	2	0.92	2
Total	182	124		
Mean L	17.3	17.6		
s	1.7	1.6		

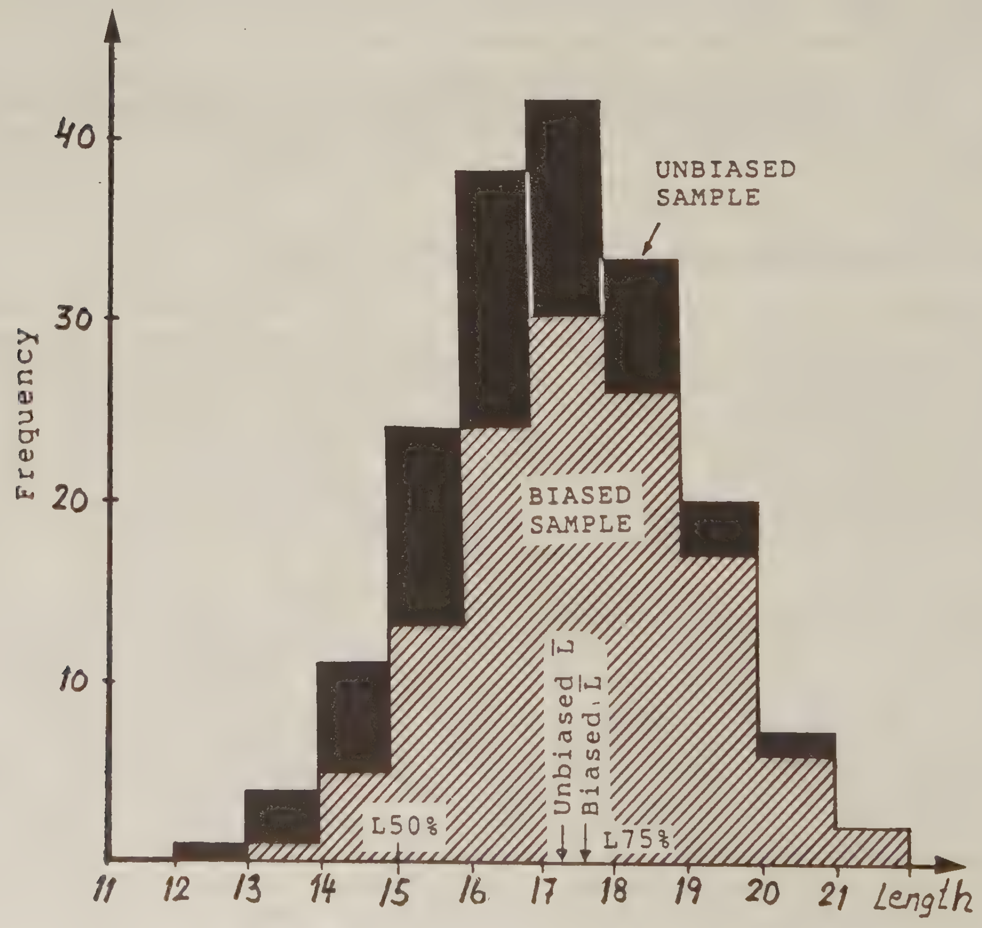


Fig. 6.5.1 Bias created by selection

In general, the effect of trawl selection is:

1. Overestimation of mean length
2. Underestimation of standard deviation
3. Underestimation of cohort size

The bias caused by selection is illustrated by Fig. 6.5.1.

Correction for mesh selection should preferably be made by means of selection curves determined by experiments like trawling with a covered codend (Section 6.1.3) or by analysis of fishing mortality arrays from cohort analysis (Section 6.4). Sometimes, selection curves for similarly shaped, closely related species can be used.

When no such data are available a selection curve may perhaps be estimated from a linearized catch curve (Section 6.1.2). Doing this, however, we encounter a logical problem because the estimation of selection is the last part of the catch curve analysis and therefore based on parameter estimates which are biased by the effects of selection. We want to start with the correction and fortunately, it is possible to do so. The sequence of analyses can be as follows:

1. Estimate L_∞ by the modified Wetherall et al. method

2. Correct length frequencies for selection using the value 1.0 for the curvature parameter, K

3. Separate normally distributed components by the Bhattacharya method (Section 3.4.1)

4. Use the estimated mean values in modal progression analysis to estimate growth parameters (Section 3.4.2)

5. Estimate Z using length converted catch curve analysis with the estimated growth parameters (Section 4.4.5).

This procedure is applicable because the estimation of a selection ogive is not very sensitive to the choice of the curvature parameter, K.

As an example we apply the modified Wetherall et al. method to the data in Table 6.1.2.1 column D (the numbers caught) and find L_∞ = 49.7 cm (cf. Table 4.5.4.1 which uses the same input data to estimate L_∞ by the Wetherall et al. method). Redoing the calculations of Table 6.1.2.1 with K = 1.0 gives:

L50% = 14.9 cm and L75% = 17.8 cm

whereas the values found with K = 0.3 were (cf. Section 6.1.2)

L50% = 15.0 cm and L75% = 18.0 cm

Using L50% = 14.9 cm and L75% = 17.8 cm gives the selection ogive (figures in brackets are from Table 6.1.2.1 for comparison):

Midlength	4	6	8	10	12	14	16	18	20	22 cm
S	.03 (.03)	.06 (.06)	.10 (.10)	.17 (.16)	.27 (.27)	.42 (.42)	.60 (.59)	.76 (.75)	.88 (.87)	.94 (.94)

Thus, the results are the same except for roundings, irrespective of the K-value.

For gill nets or any other gear with a bell-shaped selection curve one has to be careful in the interpretation of length frequency samples. The mode observed (usually there is only one mode for this type of gear) may have little to do with a cohort, but it may primarily reflect the selection curve of the gear (cf. Section 6.2).

In principle, the above mentioned method can be applied to any type of gear selection curve, but the narrower the length range selected by the gear the more difficult it is to estimate the length frequency you would have got with a non-selective gear.

(See **Exercise(s)** in Chapter 17).

7 SAMPLING COMMERCIAL CATCHES

This chapter discusses some aspects of collecting length composition data and data to estimate the weight of the total catch. For cohort analysis, only the length composition in numbers is required. However, the estimation of the weight of the total catch is closely connected to the estimation of the length composition so it is convenient to discuss the two objectives together. One basic problem in data sampling is that in principle, we want a random sample. The individual landing, however, is seldom a random sample of the population because it reflects for instance the locality fished and the properties of the vessel. Moreover, the landings are often sorted into market categories as for instance fish for direct human consumption and fish for industrial purposes. Therefore fish of the same species may occur in more than one market category, and it requires careful planning to estimate the total amount and the length composition of a given species.

It is not possible to give a complete set of general rules of how best to collect data. Each fishery has its own characteristics, which must be taken into account when planning a sampling scheme for it. However, some generalisation is possible. First of all, it must be decided which species should be investigated. This decision, of course, must be based on the economic importance of the various species (or species groups) and the funds and personnel available for the sampling. There may also be other considerations, e.g. we may want to investigate the potentials of a species which we believe is underexploited, and which could perhaps be developed into a fishery of economic importance. However, it is suggested as a rule of thumb to limit data on collection to the economically most important species. Such data also include data on juveniles landed as "trash fish", although they have little value as such.

7.1 An example of data collection

Now, we assume that we have decided on a number of species to be sampled, and let us consider the sampling of data for one of them. Instead of trying to develop a set of general rules, an (hypothetical) example is discussed. This example is believed to have many of the main features of a demersal tropical fishery. The numerical quantities of the example are arbitrary ones. To avoid too many calculations and too large tables we consider a small fishery. However, it should be relatively easy to generalize from this example and to modify it to account for special problems. The example deals with the estimation of the total catch in weight and the length composition of one species "s" caught in a mixed fishery. The catch in weight of course must refer to a prespecified period, say, a month or a quarter of the year. It is assumed that all fish caught originate from one stock and that all landings from this stock are monitored in the sampling scheme.

The physical features of the example are shown in Fig. 7.1.1. We assume the stock confined to one fishing ground exploited only by the boats from three landing places (labeled, I, II and III in Fig. 7.1.1). There are three different types of boats (labeled a, b and c). Each landing place is the home port to a number of each boat type. As an example, the three boat types could be:

a: large trawler
b: gill netters
c: "baby"-trawlers

A batch of similar boats is called a fleet. Because locations of landing places are factors of practical importance in a sampling scheme, each fleet is divided into three parts according to landing place, so that we

Table 7.1.1 The procedure for estimating total catch in weight for one species, s, during one period, t. (Compare Fig. 7.1.1)

TOTAL CATCHES OF SPECIES "s" DURING TIME PERIOD "t" IN WEIGHT UNITS								
indices			effort	observed catch	average CPUE =	estimated catch CPUE * E	total catch by landing place	total catch by fleet
time period t	fleet f	landing place h	E(t,f,h)	Wt(s,t,f,h)	Σ Wt/Σ E	Wt(s,t,f,h)	Wt(s,t,f,h)	Wt(s,t,f,*)
t	1 (a)	1 (I)	15	57	$\frac{57+83}{15+20} = 4$	-	57	172
t	1 (a)	2 (II)	20	83		-	83	
t	1 (a)	3 (III)	8	-		4* 8 = 32	32	
t	2 (b)	1 (I)	25	55	$\frac{55+105}{25+55} = 2$	-	55	184
t	2 (b)	2 (II)	55	105		-	105	
t	2 (b)	3 (III)	12	-		2*12 = 24	24	
t	3 (c)	1 (I)	20	-	$\frac{25}{25} = 1$	1*20 = 20	20	75
t	3 (c)	2 (II)	25	25		-	25	
t	3 (c)	3 (III)	30	-		1*30 = 30	30	
						grand total: Wt(s,t,*,*) =	431	

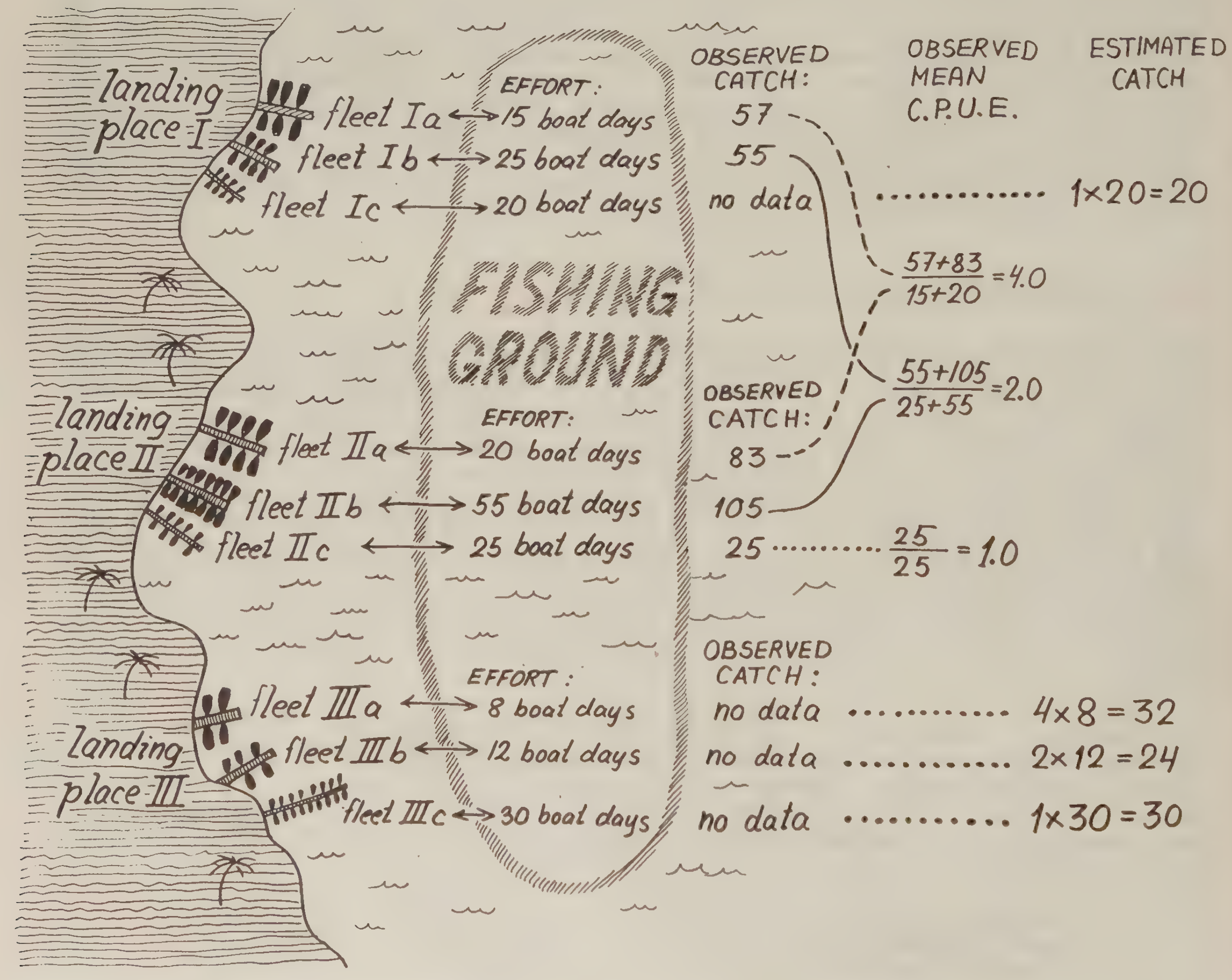

landing place I
fleet Ia ↔ 15 boat days
fleet Ib ↔ 25 boat days
fleet Ic ↔ 20 boat days
EFFORT:
FISHING GROUND
OBSERVED CATCH:
57
55
no data
OBSERVED MEAN C.P.U.E.
ESTIMATED CATCH
1×20=20
57+83 / 15+20 = 4.0
55+105 / 25+55 = 2.0
landing place II
fleet IIa ↔ 20 boat days
fleet IIb ↔ 55 boat days
fleet IIc ↔ 25 boat days
EFFORT:
OBSERVED CATCH:
83
105
25
25 / 25 = 1.0
landing place III
fleet IIIa ↔ 8 boat days
fleet IIIb ↔ 12 boat days
fleet IIIc ↔ 30 boat days
EFFORT:
OBSERVED CATCH:
no data
no data
no data
4×8=32
2×12=24
1×30=30

Fig. 7.1.1 Hypothetical example to illustrate the estimation of the total catch in weight during a time period, say a quarter of a year

operate with nine fleets in the case of Fig. 7.1.1.

We assume that effort has been observed for each boat type in each landing place. In the present example effort has been measured as the number of boat days during the period of time considered. However, a number of alternative measures of effort could have been used. In the worst case only the number of boats in each fleet is known. In that case one must assume an average number of effort units per boat per time unit. If we know, for example, the average number of boat days per month for one landing place, this figure can be applied to landing places where this information is not available.

The number of effort units by fleet no. f from landing place no. h during time period no. t (e.g. quarter no. t) is denoted:

$$E(t,f,h) = \text{effort}$$

The values of f and h can be seen in Table 7.1.1 and Fig. 7.1.1. For the moment we confine the description of the sampling scheme to a single time period, so for the moment the index "t" is a constant, e.g. second quarter of 1978. Later, data from different time periods are combined and then t becomes a variable index. The observed (or estimated) values of the effort E(t,f,h) for the example are shown in Table 7.1.1 and Fig. 7.1.1.

7.2 Estimation of total catch in weight

Assume that due to shortage of funds and personnel the total catches have not been recorded for any of the fleets of landing place III or for fleet c of landing place I. The observed total catches for the remaining components of the fishery are shown in Table 7.1.1 (and Fig. 7.1.1). We imagine these data to originate from the files of the fishermen's cooperative or some other body. They may also have been collected by the staff of the fisheries research institute. We denote by

$$Wt(s,t,f,h)$$

the total weight of species s landed during period t by fleet f to landing place h.

From those combinations of fleet and landing place where effort and catch data are available, average catch per unit of effort, CPUE, can be calculated (cf. Table 7.1.1 and Fig. 7.1.1). CPUE here stands for catch in weight per unit of effort, whereas the CPUE concept we used in connection with estimation of Z was in units of number of fish per unit of effort (cf. Section 4.3). These CPUE figures can be used to estimate catch for those fleets for which only effort data were available:

$$\text{estimated catch} = \text{effort} * \text{CPUE}$$

as demonstrated in Table 7.1.1 and Fig. 7.1.1

We are now in a position to calculate the total catch of species s during time period t for each fleet. To avoid too many summation signs we introduce a more convenient notation. Whenever an index is replaced by "*" it means the sum over the index in question. For example:

$$Wt(s,t,f,*) = \sum_{h=1}^{3} Wt(s,t,f,h)$$

is the catch by fleet f from all landing places, and

Table 7.3.1 Raising the catch of the sampled trips to the total catch of one fleet in one landing place (compare Figs. 7.3.1 and 7.3.4)

length group	total sampled (from Fig. 7.3.4) Cs(s,t,f,h,*,i)	raised to total catch Cr(s,t,f,h,*,i)
5-6	51	191.25
6-7	148	555
7-8	216	810
8-9	93	348.75
9-10	63	236.25
10-11	176	660
11-12	122	457.5
12-13	58	217.5
13-14	15	56.25
total	942 = Cs(s,t,f,h,*,*)	3532.5 = Cr(s,t,f,h,*,*)

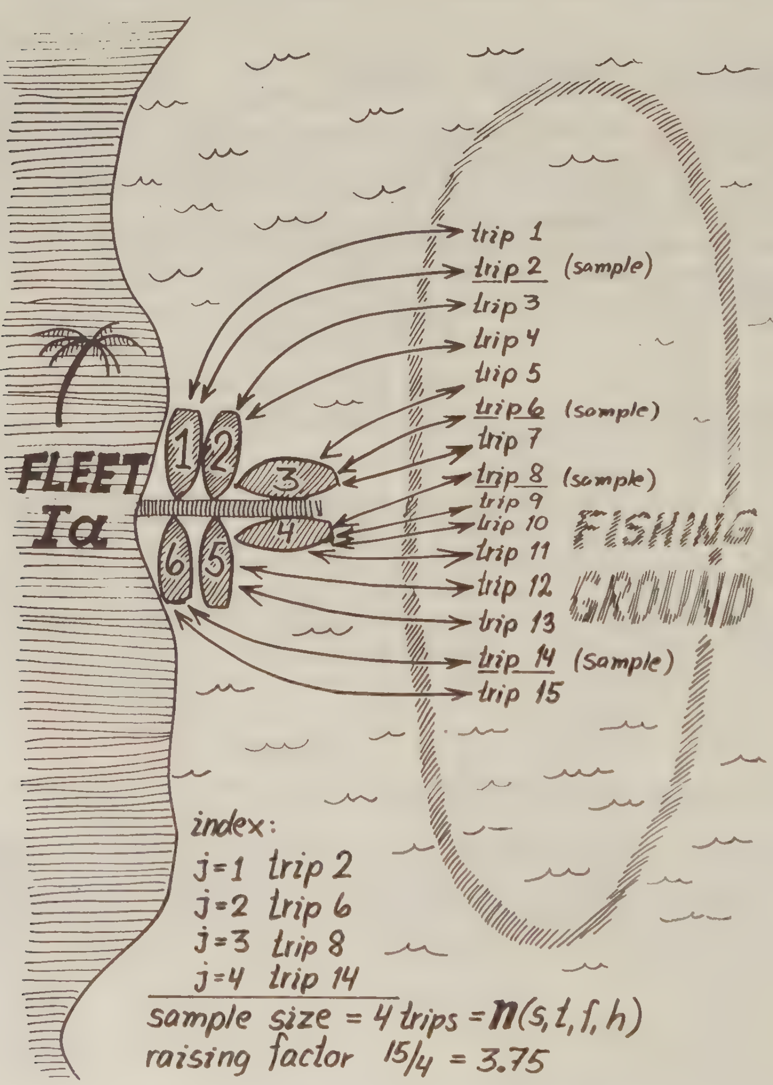

FLEET Ia
1
2
3
4
5
6
trip 1
trip 2 (sample)
trip 3
trip 4
trip 5
trip 6 (sample)
trip 7
trip 8 (sample)
trip 9
trip 10
trip 11
trip 12
trip 13
trip 14 (sample)
trip 15
FISHING GROUND
index:
j=1 trip 2
j=2 trip 6
j=3 trip 8
j=4 trip 14
sample size = 4 trips = n(s,t,f,h)
raising factor 15/4 = 3.75

Fig. 7.3.1 Illustration of sampling for the estimation of the length composition of the catch of one fleet from one landing place

$$Wt(s,t,*,*) = \sum_{f=1}^{3} \sum_{h=1}^{3} Wt(s,t,f,h) = \sum_{f=1}^{3} Wt(s,t,f,*)$$

is the total catch of species s during time period t landed by all fleets to all landing places (cf. Table 7.1.1). So much for the total catch of species s in weight units.

We now turn to the main objective, the derivation of the length composition of species s, i.e. the derivation of input to the cohort analysis.

7.3 Estimation of total length composition

The basic sample for estimation of length compositions is associated with a "trip" in the present example. Each fleet from each landing place makes a number of trips during the time period considered. In Fig. 7.3.1 fleet Ia is considered as an example. During this period 15 trips were carried out by fleet Ia. At the end of some of the trips samples were collected on the jetty when unloading the catch. Usually, it is impossible to take samples from all trips. In the example of Fig. 7.3.1 four out of the 15 trips were sampled. (These should be selected at random, see the discussion in Section 2.7). Let the number of samples be denoted:

$n(t,f,h)$

In the example of Fig. 7.3.1: $n(t,1,1,) = 4$. Each sample is given a number, or an index, j, for our internal book-keeping:

$j = 1,2,\ldots,n(t,f,h)$ (cf. Fig. 7.3.1)

The situation when sampling a trip is illustrated in Fig. 7.3.2. The catch of a trip is assumed to consist of two categories

1. FISH FOR DIRECT HUMAN CONSUMPTION or "CONSUMPTION" FISH. This part is sorted into species (or species groups) and contains the marketable sizes

2. "TRASH FISH". This part is not sorted by the fishermen. It contains fish species not used for human consumption including valuable fish species below the marketable size.

Samples are taken from both categories, since from a biological point of view they are of equal importance.

SAMPLING THE CATCH FOR HUMAN CONSUMPTION FROM ONE TRIP:

A sample is taken from each of the more important species of the consumption fish and the length of each fish is recorded. Let i be the index of a length group. We denote by

$x(s,t,f,h,j,i)$

the number of fish of species s in sample no. j belonging to:

length group	i	landing place	h
fleet	f	time period	t

According to the notation introduced in Section 7.2 the total number of fish in the sample is:

$x(s,t,f,h,j,*)$

(see example in the middle part of Fig. 7.3.3 in which case $x(s,t,f,h,j,*) = 37$).

The weight of the sample is also recorded, and we use the symbol:

$ws(s,t,f,h,j)$ = weight of the $x(s,t,f,h,j,*)$ fish in sample j (see Fig. 7.3.2)

The weight of the total consumption catch of species s is recorded:

$Ws(s,t,f,h,j)$ = weight of consumption catch of species s.

Thus, the small letters refer to the sample and capital letters to the total catch. The suffix "s" stands for "sampled". In the example of Fig. 7.3.2 a sample of

$$4.7 \text{ kg} = ws(s,t,f,h,j)$$

was taken from at total catch of

$$14.1 \text{ kg} = Ws(s,t,f,h,j).$$

The estimate of the total number of consumption fish of species s in length group i becomes:

$$x(s,t,f,h,j,i) \frac{Ws(s,t,f,h,j)}{ws(s,t,f,h,j)}$$

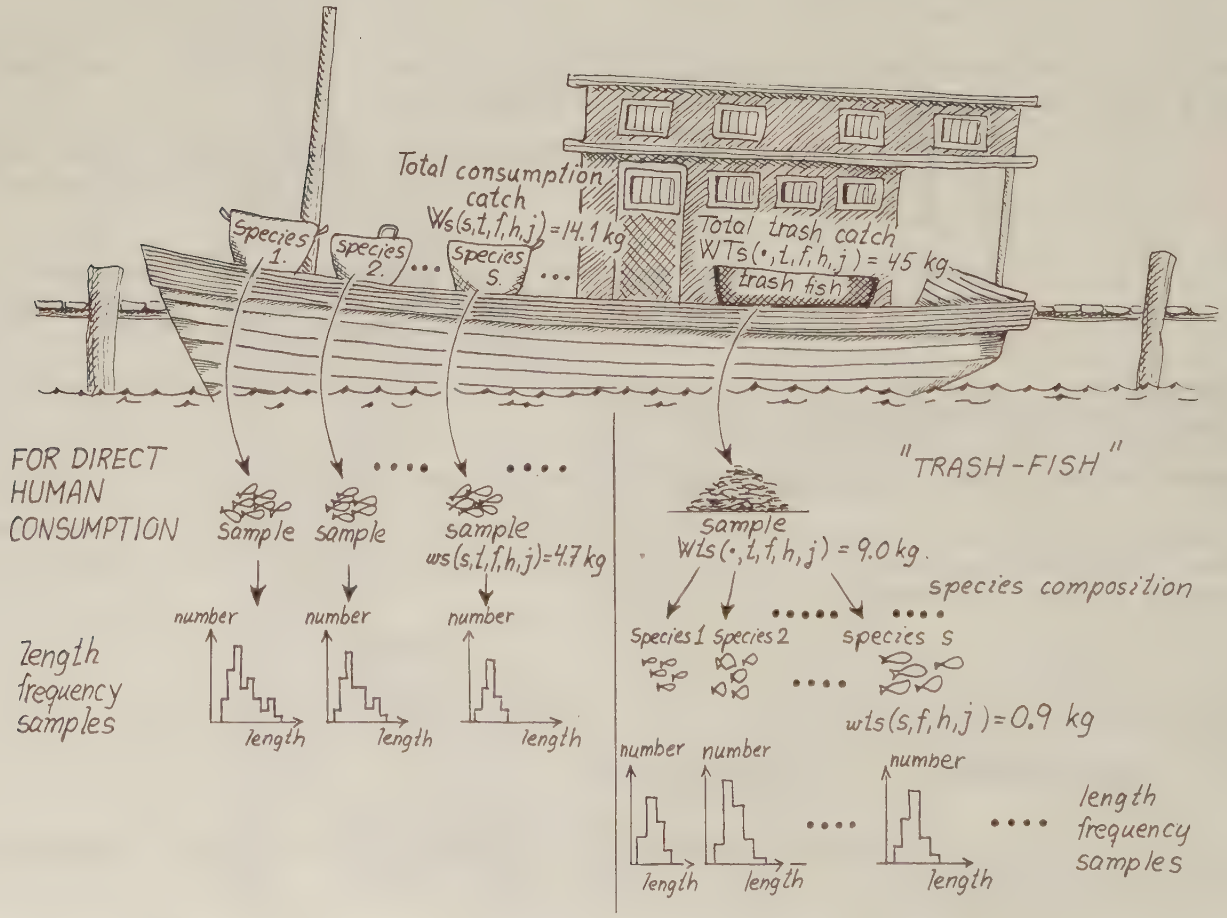

Total consumption catch
Ws(s,t,f,h,j) = 14.1 kg
Species 1.
species 2.
...
Species S.
...
Total trash catch
WTs(•,t,f,h,j) = 45 kg
trash fish
FOR DIRECT HUMAN CONSUMPTION
sample
sample
sample
ws(s,t,f,h,j)=4.7 kg
number
length
length frequency samples
"TRASH-FISH"
sample
Wts(•,t,f,h,j) = 9.0 kg
species composition
Species 1
Species 2
species s
wts(s,f,h,j) = 0.9 kg
number
length
length frequency samples

Fig. 7.3.2 Sampling from a single trip (for further explanation, see text)

We say that the sample has been raised to the total by the raising factor Ws/ws. In the example the raising factor is (see middle part of Fig. 7.3.3.): $Ws/ws = 14.1/4.7 = 3.0$

SAMPLING THE TRASH FISH FROM ONE TRIP:

Let the total weight of the trash fish be

$WTs(*,t,f,h,j)$ kg

(index "s" is replaced by "*" because the trash fish is not sorted for species). From the WTs kg is taken a sample of

$wts(*,t,f,h,j)$ kg

This sample is then separated into groups containing a group for each of the consumption species and a number of other abundant species plus a remaining group of rare, low valued species (see Fig. 7.3.2). The weight of species s fish from the trash sample is

$wts(s,t,f,h,j)$ kg

These fish are measured and the lengths recorded. The number of length group i fish is

$xt(s,t,f,h,j,i)$

(see upper part of Fig. 7.3.3). They are raised to account for the total trash catch by the raising factor WTs/wts:

$$xt(s,t,f,h,j,i)\,\frac{WTs(*,t,f,h,j)}{wts(*,t,f,h,j)}$$

In the example of Fig. 7.3.2 $Wts(*,t,f,h,j) = 45$ kg and the sample weight is $wts(*,t,f,h,j) = 9$ kg so that the raising factor becomes $WTs/wts = 45/9 = 5.0$.

COMBINING CONSUMPTION FISH SAMPLE AND TRASH FISH SAMPLE FROM ONE TRIP:

An estimate of the total number by length group, c, is obtained by simple addition of consumption and trash fish estimates:

$$Cs(s,t,f,h,j,i) = x(s,t,f,h,j,i)\,\frac{Ws}{ws} + xt(s,t,f,h,j,i)\,\frac{WTs}{wts}$$

(see lower part of Fig. 7.3.3). For example, the estimated total number for length group 8-9 cm becomes:

$$2 * \frac{14.1}{4.7} + 3 * \frac{45}{9} = 2 * 3 + 3 * 5 = 6 + 15 = 21$$

ADDING SAMPLES FROM SEVERAL TRIPS:

Again a simple summation will do. The estimated number caught by length group in all n(t,f,h) samples is:

$$Cs(s,t,h,*,i) = \sum_{j=1}^{n(t,f,h)} Cs(s,t,h,j,i)$$

(see example in Fig. 7.3.4.). For example, the estimated total number of all four sampled trips for length group 8-9 cm becomes:

$$25 + 22 + 20 + 26 = 93$$

Trash-fish :

Total weight of catch: $WT_s = 45$ kg

weight of sample : $wt_s = 9$ kg

raising factor: $\frac{WT_s}{wt_s} = \frac{45}{9} = 5.0$

For direct human consumption :

Total weight of catch: $W_s = 14.1$ kg

weight of sample $w_s = 4.7$ kg

raising factor: $\frac{W_s}{w_s} = \frac{14.1}{4.7} = 3.0$

Total length composition for species: s ; time period t ; fleet f ; landing place h ; sample no j (trip)

$$C_s(s,t,f,h,j,i) = xt\frac{WT_s}{wt_s} + x\frac{W_s}{w_s}$$

or with all indices :

$$xt(s,t,f,h,j,i)\frac{WT_s(s,t,f,h,j)}{wt_s(s,t,f,h,j)} + x(s,t,f,h,j,i)\frac{W_s(s,t,f,h,j)}{w_s(s,t,f,h,j)}$$

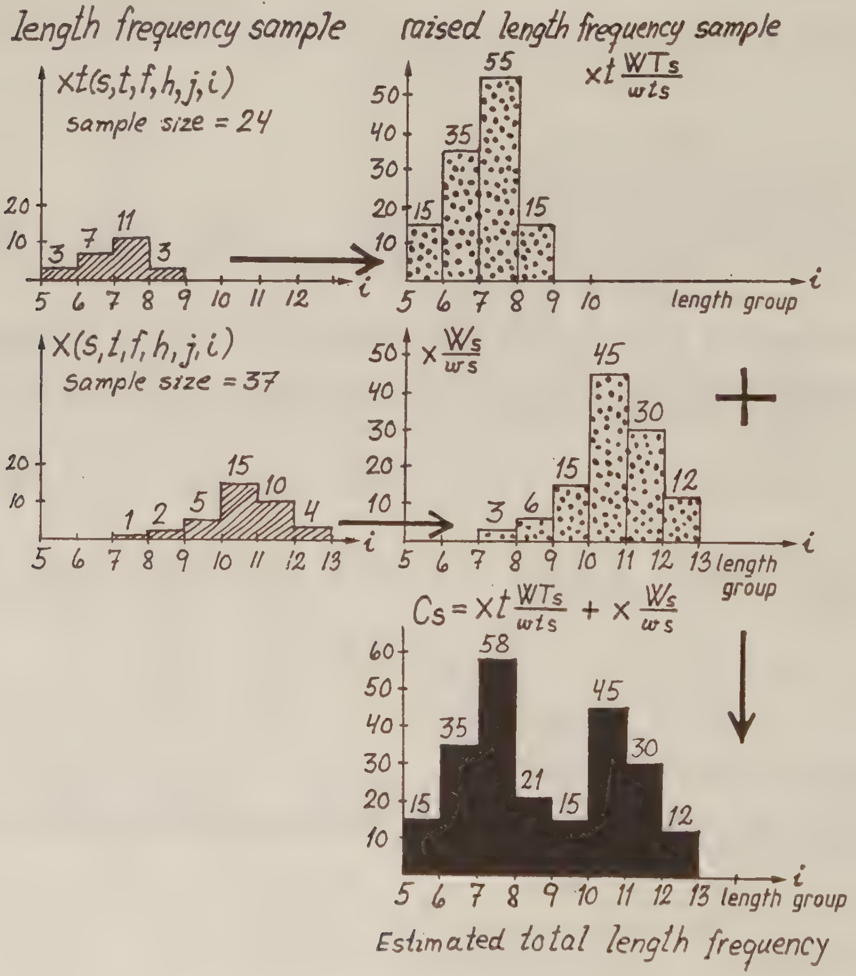


Fig. 7.3.3 Combining length samples from catch for human consumption and trash fish

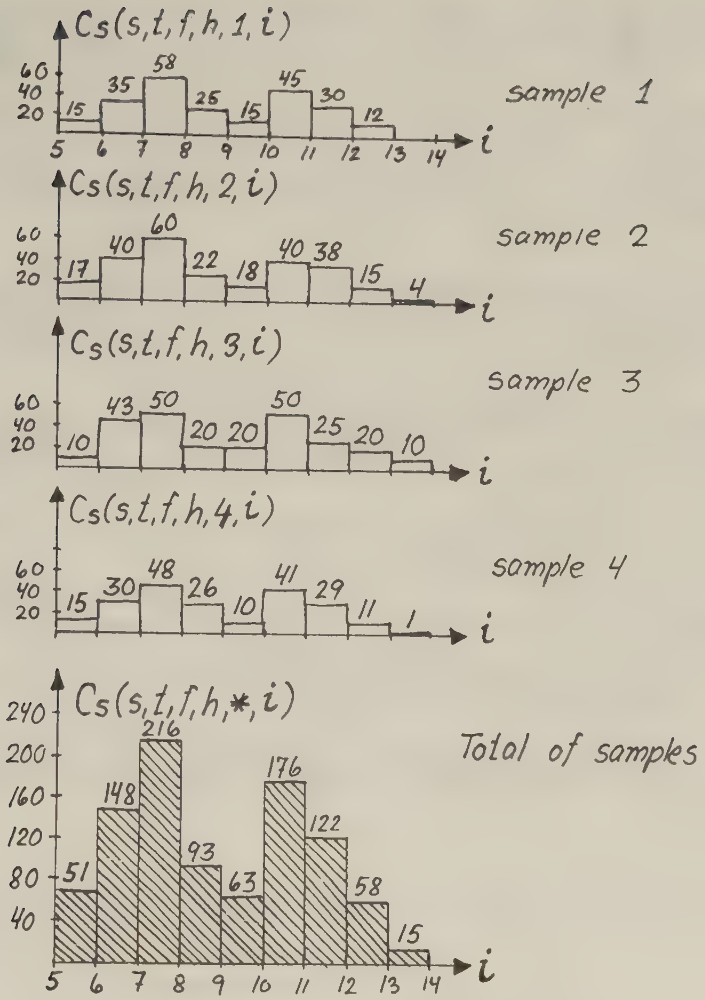


Fig. 7.3.4 **Adding length compositions of species s of sampled trips. The sample used in the example of Fig. 7.3.3. appears as sample no. 1**

RAISING THE SAMPLED TRIPS TO THE TOTAL CATCH OF THE FLEET IN ONE LANDING PLACE:

The total length distribution of the sampled trips Cs(s,t,f,h,*,i) can be raised to account for the entire catch in period t, by fleet f to landing place h, by using the raising:

$$Cr(s,t,f,h,*,i) = \frac{\text{total number of trips}}{\text{number of trips sampled}}\, Cs(s,t,f,h,*,i)$$

where the suffix r stands for "raised". In the example (Fig. 7.3.1) the raising factor is 15/4 = 3.75 and the result is shown in Table 7.3.1.

This raising procedure is reasonable only if a "trip" is a well-defined unit of effort. If some trips are of a duration of, say, one fishing day and others are of a duration of 5 fishing days, a "fishing day" might be a more appropriate unit. Either unit, "trip" or "fishing day" makes sense only if the boats of a fleet are fairly similar, i.e. have the same fishing power. Other possible units of effort are "number of man days", "number of trawling hours", "number of gill net sets", etc.

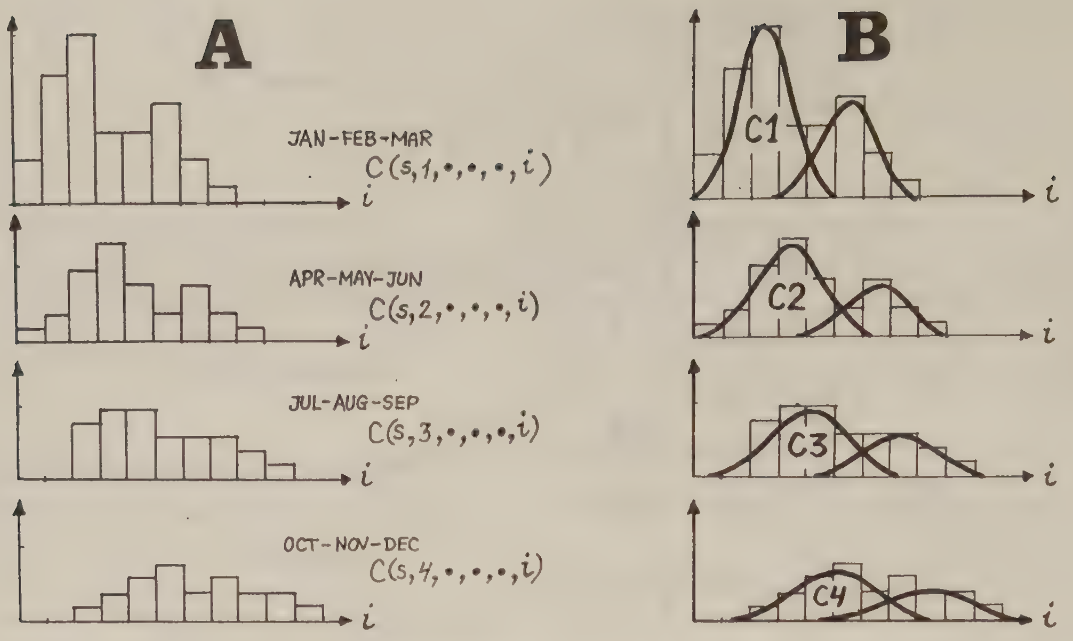


Fig. 7.3.5 A: Total length frequencies by quarter year.

B: Length frequencies resolved into normally distributed cohort components (input to cohort analysis)

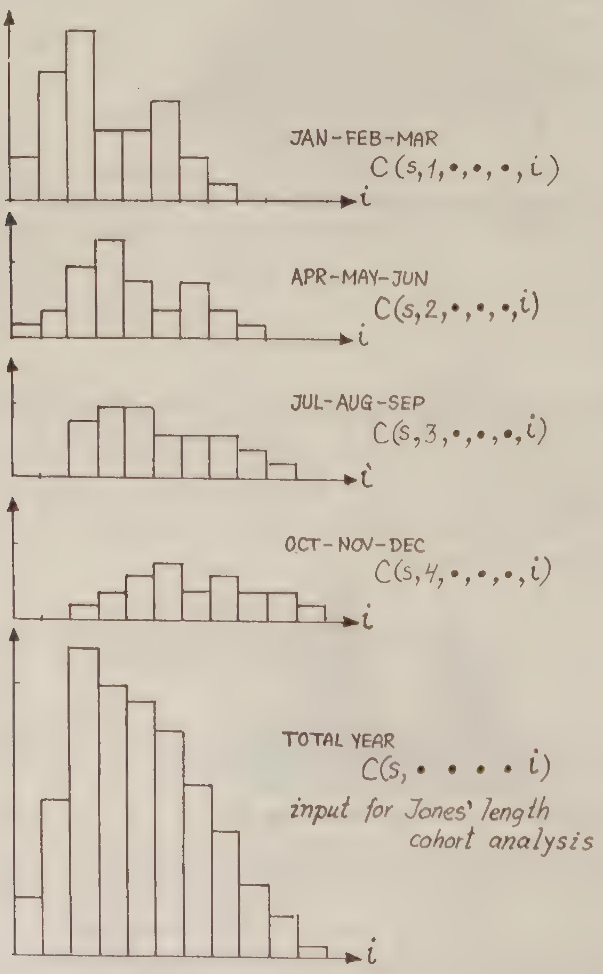


Fig. 7.3.6 Summing total length compositions of time periods (total length composition is input for Jones' cohort analysis)

ADDING SAMPLED LANDING PLACES FOR ONE FLEET AND RAISING TO ALL LANDING PLACES:

The total length distribution of all sampled landing places (cf. Fig. 7.1.1) is obtained by simple addition:

$$Cr(s,t,f,*,*,i) = \sum_{h} Cr(s,t,f,h,*,i)$$

This figure can be raised to the total for all landing places by applying the raising:

$$C(s,t,f,*,*,i) = \frac{\text{total effort of all landing places}}{\text{effort of sampled landing places}} \; Cr(s,t,f,*,*,i)$$

In the example of Fig. 7.1.1 and Table 7.1.1 the raising factors are:

fleet	total effort	effort of sampled landing places	raising factor
1 (a)	43	35	43/35 = 1.229
2 (b)	92	80	92/80 = 1.150
3 (c)	75	25	75/25 = 3.000

ADDING FLEETS:

This too is a simple summation. The total length distribution of species s caught during time period t is:

$$C(s,t,*,*,*,i) = \sum_{f} C(s,t,f,*,*,i)$$

Depending on the type of analysis required we may stop the processing of catch data at this level. The result in that case is the four quarterly length frequencies shown in Fig. 7.3.5A. These may be resolved into cohort components by, for example, the Bhattacharya method (Section 3.4) as shown in Fig. 7.3.5B, and cohort analysis can be applied. The number caught from a cohort (e.g. the numbers C1,C2,C3 and C4 indicated in Fig. 7.3.5B) form the input to Pope's cohort analysis, the procedure of Table 5.2.1 using Eqs. 5.2.3 and 5.2.4. A quarter of a year is now the time unit (from t to t+1) in Fig. 5.2.1.

Alternatively, we may choose to apply Jones' length cohort analysis. In this case we do not separate the cohorts, but proceed as follows:

ADDING TIME PERIODS:

This is the final step, which gives the length frequency for the year. The adding is simple:

$$C(s,*,*,*,*,i) = \sum_{t} C(s,t,*,*,*,i)$$

Fig. 7.3.6 shows an example in which quarterly length compositions are summed to an annual length composition. These C-values can be used as input to length cohort analysis. We may also use the average values for a range of years (see Section 5.2).

The resolution into normally distributed components shown in Fig. 7.3.5 becomes more problematic the longer the sampling period is. The quarterly samples show a certain cohort structure, whereas the cohorts are disguised in the length frequency distribution for the whole year. This example illustrates that for age based cohort analysis we must work with relatively short time periods. Otherwise we are unable to identify the cohorts. We are not interested in the slopes of length frequencies in this case, only in the numbers caught from each cohort.

For a Jones' length cohort analysis, on the other hand, the period must be relatively long, because in this case we want the cohort structure to be leveled out. The slopes of the normally distributed cohort components are not related (in a simple way) to mortality, whereas the slope of the combined frequency is.

Finally, it is emphasized that the procedure explained above is an example which may or may not fit to your fishery. Especially the definition of the effort unit ("trip") may be inappropriate for some fisheries. Also the assumption made that the samples are taken from the boats at the time the catch is unloaded may not fit to many cases.

The LFSA package of micro computer programs (Sparre, 1987) contains a suite of programs for the above mentioned and similar data manipulations, i.e. various types of summations and raising procedures for length frequency samples.

8 THE YIELD PER RECRUIT MODEL OF BEVERTON AND HOLT

The stock and yield assessment models dealt with so far (Chapter 5) are variable parameter models aiming at describing changes in the stock and the yield from one year to another as a result of changing the fishing pattern: the effort exerted on individual age or length groups. "The yield per recruit model" (Beverton and Holt, 1957) is in principle a steady state model, i.e., a model describing the state of the stock and the yield in a situation when the fishing pattern has been the same for such a long time that all fish alive have been exposed to it since they recruited. There are extensions to the model dealing with the transitory phase after a change in the fishing pattern, but these are seldom used because models of the Thompson and Bell type (Chapter 5) provide a simpler description of non-steady-state situations.

There are some rigorous assumptions underlying the Beverton and Holt approach:

1. Recruitment is constant, yet not specified. Hence the expression "yield per recruit"

2. All fish of a cohort are hatched on the same date

3. Recruitment and selection are "knife-edge" (see Chapter 6)

4. The fishing and natural mortalities are constant from the entry to the exploited phase

5. There is a complete mixing within the stock

Again, there are extensions to the model which release some of these assumptions and again, the extensions are not particularly attractive to handle.

Fig. 8.0.1 shows the result of a stock and yield assessment with the yield per recruit model. The age at entry to the exploited phase is kept constant. The independent variable is the effort as expressed by the coefficient F of fishing mortality. The dependent variables are the exploited biomass in grammes per recruit and the annual yield in grammes per recruit. Only when the total annual yield is known in a steady state situation, for a given value of F, can the recruitment number be calculated by dividing the total yield by the yield in grammes per recruit.

The "yield per recruit curve" often has a maximum: the "maximum sustainable yield" for the mesh size (age at entry) used. "Sustainable", because higher yields can be obtained only temporarily, by a sudden increase of effort. The "biomass per recruit curve" is always decreasing with increasing effort. The curve is proportional to the catch per unit of effort, under the assumptions underlying the model. An important aspect of fish stock management is therefore clearly illustrated by the model: overfishing (the so-called "growth overfishing") occurs when the effort is so high that the total yield decreases with increasing effort, whereas a decrease in catch per unit effort (biomass) is only a sign of increased fishing, not a sign of overfishing.

Changing the age at entry to the exploited phase leads to a family of curves like Fig. 8.0.2, each with a different maximum sustainable yield. Plotting the maxima of such curves shows which combination of effort and age at entry gives the overall maximum sustainable yield.

One serious drawback of the yield per recruit model is that it disregards "recruitment overfishing" (see Section 8.3 and Chapter 12). With an ever decreasing biomass for increasing effort a point may be reached when the stock is no longer big enough to sustain the constant recruitment assumed. Yet, the model gives no warning about this.

In the section below we shall go into more detail with the aspects of the model and describe its mathematics. It is originally an age based model, but age is easily converted into length, again using the principles of the conversion of the catch curve (Section 4.4.2).

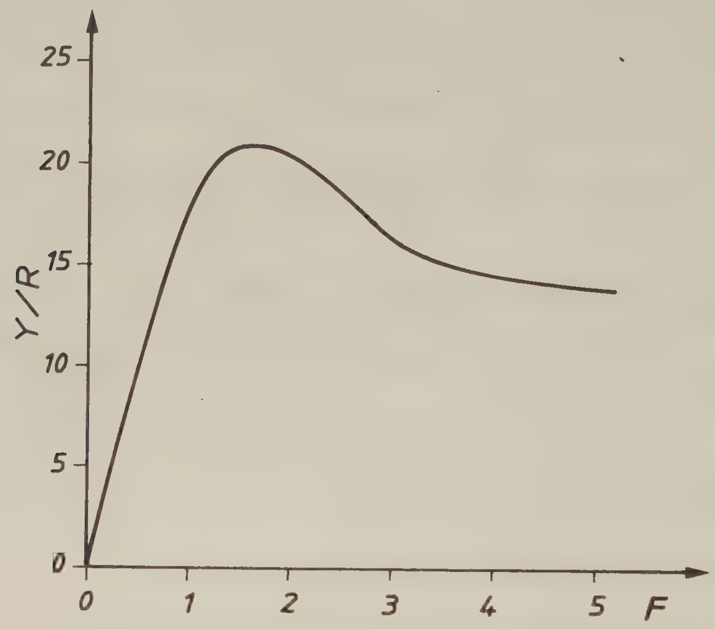


Fig. 8.0.1 Result of a stock and yield assessment with the yield per recruit model

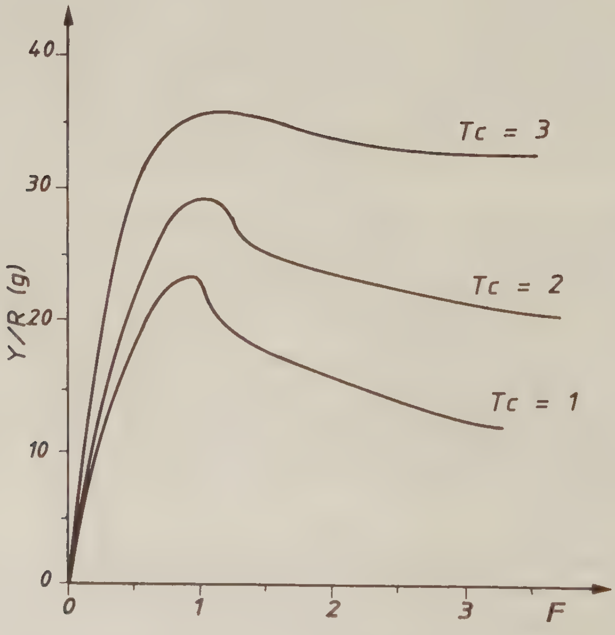


Fig. 8.0.2 Yield per recruit curves with different ages of first capture (Tc)

8.1 The exponential decay model behind the Beverton and Holt model

The assumed life history of a cohort in the Beverton and Holt model is described in Fig. 8.1.1. At age Tr the cohort recruits to the fishing ground (cf. Section 4.1), all at the same time: "knife-edge recruitment". From age Tr to age Tc it is not exposed to fishing mortality. (We assume that all fish of ages between Tr and Tc escape through the meshes if they enter the gear). Thus, in that period they suffer only from natural mortality, M, which is assumed to remain constant throughout the lifespan of the cohort. At age Tc, the "age at first capture", the cohort is assumed to be suddenly exposed to full fishing mortality, F, which is assumed to remain constant for the rest of the cohort's life. The sigmoid shaped gear selection curve introduced in Chapter 6 is approximated by the so-called "knife-edge selection" (see Fig. 8.1.1). The catch from the cohort is therefore assumed to be zero before the cohort has attained the age Tc. The number of survivors at age Tr is the recruitment to the fishery:

$$R = N(Tr) \qquad (8.1.1)$$

The number of survivors at age Tc is:

$$N(Tc) = R \exp[-M(Tc-Tr)] \qquad (8.1.2)$$

The number of survivors at age t, where t > Tc is

$$N(t) = N(Tc) \exp[-(M+F)(t-Tc)] = R \exp[-M(Tc-Tr) - (M+F)(t-Tc)] .$$

The fraction of R surviving until age t thus becomes:

$$N(t)/R = \exp[-M(Tc-Tr) - (M+F)(t-Tc)] \qquad (8.1.3)$$

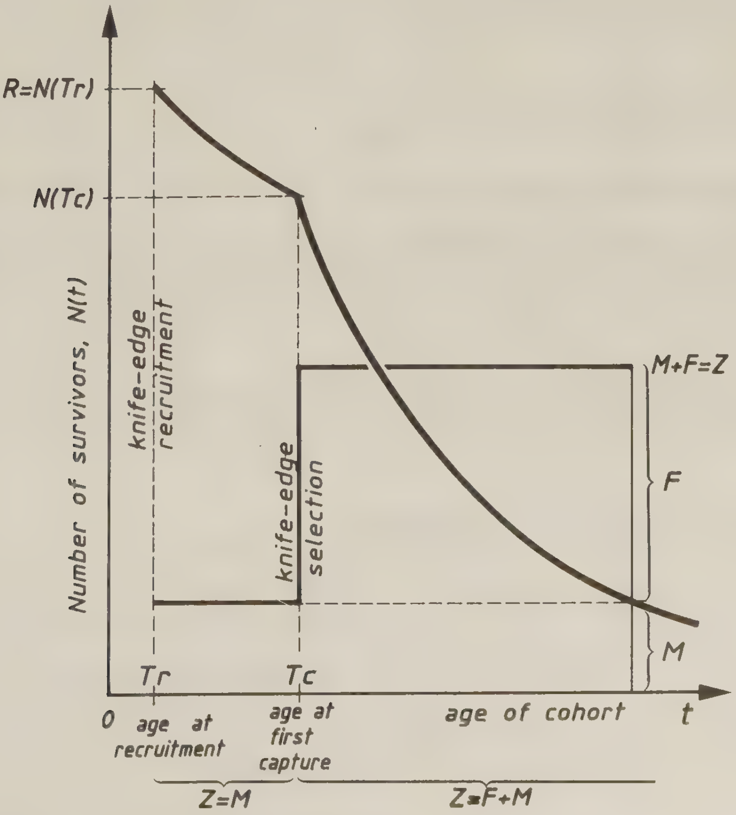

R=N(Tr)
N(Tc)
Number of survivors, N(t)
knife-edge recruitment
knife-edge selection
M+F=Z
F
M
Tr
Tc
0
age at recruitment
age at first capture
age of cohort
t
Z=M
Z=F+M

Fig. 8.1.1 The life history of a cohort as assumed in the Beverton and Holt model

8.2 Yield per recruit

Because Beverton and Holt express yields on a "per recruit basis", the yields are relative, i.e. relative to the recruitment. If, say, a recruitment of 100 million fish gives a yield of 10,000 tonnes then 200 million recruits would yield 20,000 tonnes according to the model. This assumption may appear trivial, but is not, as one could well imagine that the more abundant a species becomes the worse the conditions for the individuals will be due to, for example, food competition and cannibalism. The results of the model are expressed in units of yield per recruit (grammes per recruit). In the example above the yield per recruit becomes:

$$\frac{10\ 000\ 000\ 000}{100\ 000\ 000} = 100 \text{ grammes per recruit}$$

To derive the mathematical expression for the model we take (as usual) a starting point in the catch equation in the form of Eq. 4.2.10:

$$C(t,t+\Delta t) = \Delta t \; F \; N(t) \tag{8.2.1}$$

Eq. 8.2.1 gives the number of fish caught in the time period from t to $t+\Delta t$ from a cohort, when Δt is small. To get the corresponding yield in weight, this number should be multiplied by the individual weight of a fish. If Δt is small, then the body weight of a fish will remain approximately constant during the time period from t to $t+\Delta t$, and the yield becomes:

$$Y(t,t+\Delta t) = \Delta t \; F \; N(t) \; w(t) \tag{8.2.2}$$

where $w(t)$ is the body weight of a t years old fish, as defined by the weight converted von Bertalanffy equation (Eq. 3.3.6.1). To get the yield per recruit for the time period from t to $t+\Delta t$ Eq. 8.2.2 is divided by the number of recruits, R:

$$\frac{Y(t,t+\Delta t)}{R} = F \; \frac{N(t)}{R} \; w(t) \; \Delta t \tag{8.2.3}$$

where $N(t)/R$ is defined by Eq. 8.1.3.

Eq. 8.2.3 is the "Beverton and Holt model for a short time period". To get the total yield per recruit for the entire life span of the cohort, Y/R, all the small contributions defined by Eq. 8.2.3 must be summed:

$$\begin{aligned} Y/R = {} & Y(Tc\ ,\ Tc+\Delta t)/R + Y(Tc+\Delta t\ ,\ Tc+2\Delta t)/R + \\ & Y(Tc+2\Delta t, Tc+3\Delta t)/R + Y(Tc+3\Delta t\ ,\ Tc+4\Delta t)/R + \\ & Y(Tc+4\Delta t, Tc+5\Delta t)/R + Y(Tc+5\Delta t\ ,\ Tc+6\Delta t)/R + \\ & \dots \\ & \dots \\ & + Y(Tc+(n-1)\Delta t, Tc+n\Delta t)/R \end{aligned}$$

where "n" is some large number, so large that the number of fish older than $Tc+n\Delta t$, i.e., $N(Tc+n\Delta t)$, is so small that it can be ignored.

The next step is to convert the sum written above into a form which can easily be calculated (e.g. by using a pocket calculator). If the number of terms, n, in the sum is large (and it must be large to make the approximation for $w(t)$ a reasonable one) it will take a long time to do this summing. However, by using a long series of mathematical derivations which is outside the scope of this manual, one can show that the sum above can be written in a more convenient way:

$$Y/R = F \exp[-M(Tc-Tr)]\ W_\infty \left[\frac{1}{Z} - \frac{3S}{Z+K} + \frac{3S^2}{Z+2K} - \frac{S^3}{Z+3K} \right] \qquad (8.2.4)$$

where:

$S = \exp[-K(Tc-t_o)]$
K = von Bertalanffy growth parameter
t_o = von Bertalanffy growth parameter
Tc = age at first capture
Tr = age at recruitment
W_∞ = asymptotic body weight
F = fishing mortality
M = natural mortality
$Z = F + M$, total mortality

Eq. 8.2.4 is the "Beverton and Holt yield per recruit model" (1957), written in the form suggested by Gulland (1969). Although the equation looks complicated, it is quite easy to handle with a programmable pocket calculator.

The two parameters F and Tc are those which can be controlled by fishery managers, because:

1) F is proportional to effort (cf. Eq. 4.3.0.7)

2) Tc is a function of gear selectivity

Therefore, Y/R is considered a function of F and Tc. Most often you will see Y/R plotted against F (or effort).

EXAMPLE

As an example we calculate the Y/R for *Nemipterus marginatus* as a function of F, using the following parameters:

$K = 0.37$ per year $\quad Tc = 1.0$ year $\quad t_o = -0.2$ year
$M = 1.1$ per year $\quad Tr = 0.4$ year $\quad W_\infty = 286$ grammes

We start by calculating the terms in Eq. 8.2.4 wich are independent of F:

$$S = \exp[-K(Tc-t_o)] = \exp[-0.37 * (1.0+0.2)] = 0.6415$$

$$3S = 1.9244 \quad , \quad 3S^2 = 1.2344 \quad , \quad S^3 = 0.2639$$
$$M+K = 1.47 \quad , \quad M+2K = 1.84 \quad , \quad M+3K = 2.21$$

$$\exp[-M(Tc-Tr)]\ W_\infty = \exp[-1.1 * (1.0-0.4)] * 286 = 147.8$$

Inserting these values into Eq. 8.2.4 we get:

$$Y/R = F * 147.8 * \left[\frac{1}{F+1.1} - \frac{1.9244}{F+1.47} + \frac{1.2344}{F+1.84} - \frac{0.2639}{F+2.21} \right]$$

To produce a graph corresponding to Fig. 8.2.1 this expression must be evaluated for a range of F-values, giving a sufficient number of points to fit the curve by eye. For example for F = 0.5:

$$Y/R = 0.5 * 147.8 \left[\frac{1}{0.5+1.1} - \frac{1.9244}{0.5+1.47} + \frac{1.2344}{0.5+1.84} - \frac{0.2639}{0.5+2.21} \right]$$

$$= 0.5 * 147.8 * 0.0785 = 5.8 \text{ grammes per recruit}$$

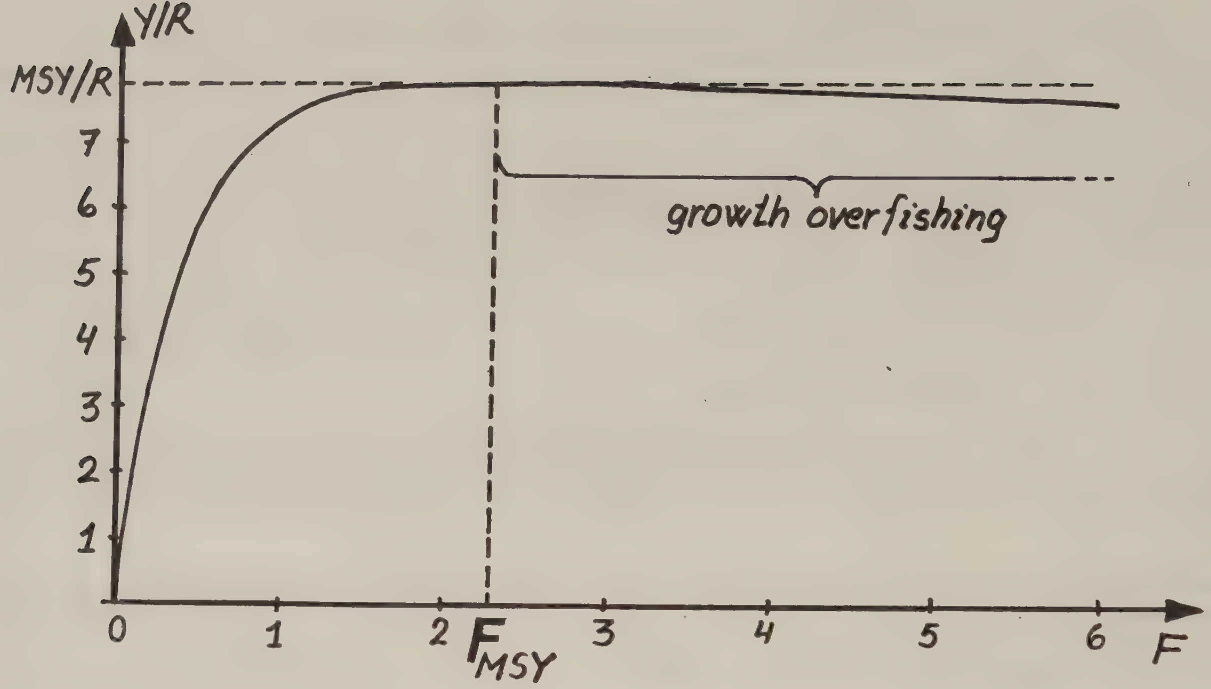


Fig. 8.2.1 Yield per recruit curve as a function of F for the parameters:

K = 0.37 per year Tc = 1.0 year t_o = -0.2 year
M = 1.1 per year Tr = 0.4 year W_∞ = 286 grammes

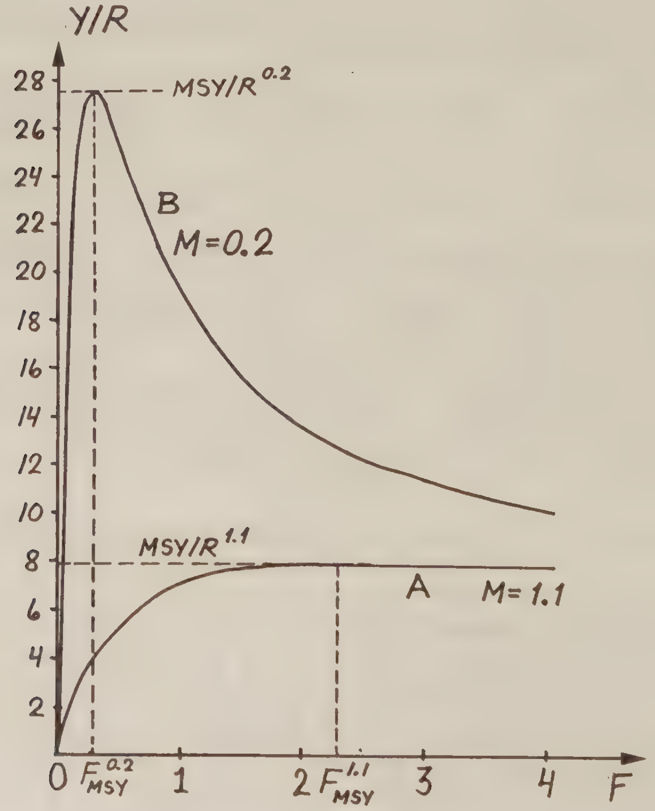


Fig. 8.2.2 Yield per recruit curve as a function of F for the parameters:

K = 0.37 per year Tc = 1.0 year t_o = -0.2 year
A: M = 1.1 per year Tr = 0.4 year W_∞ = 286 grammes
B: M = 0.2 per year

and for F = 0.6:

$$Y/R = 0.6 * 147.8 \left[\frac{1}{0.6+1.1} - \frac{1.9244}{0.6+1.47} + \frac{1.2344}{0.6+1.84} - \frac{0.2639}{0.6+2.21} \right]$$

= 6.3 grammes per recruit.

Repeating this calculation for F = 0.1, 0.2, 0.3, ... till 6.0 enables us to produce the graph shown in Fig. 8.2.1.

By testing various F-values it is found that F = 2.3 gives the maximum value of Y/R, the "Maximum Sustainable Yield per Recruit" (cf. Section 8.0):

MSY/R = 7.9 grammes per recruit

which corresponds to the optimum fishing mortality:

F(MSY) = 2.3 per year

(see Fig. 8.2.1). Because the Y/R-model assumes a constant parameter system (cf. Section 4.4.1) the results to be read from the curve only apply after the system has had constant parameters for a while. When F is changed it takes some time before the Y/R becomes the one predicted by the curve. How long this transient period is depends on the longevity of the species in question. This feature is the reason for the prefix "sustainable" to the yield. Inspecting Fig. 8.2.1 it appears that the F(MSY)-level is not determined with any great precision. Actually, for F > 1.5 the yield remains the same for a wide range of effort.

In Fig. 8.2.2, curve B, is a Y/R-curve with a different shape compared to that of Fig. 8.2.1 (which is reproduced as curve A). Curve B has a pronounced maximum, it has a lower value of F(MSY) and a higher value of MSY/R compared to curve A. The only difference between the two curves is the value of the natural mortality rate, M viz. as M = 0.2 per year for curve B and M = 1.1 per year for curve A. The lessons learned from Fig. 8.2.2 are general ones:

1. A lower M produces a lower F(MSY) and a higher MSY/R. Fishing effort levels above F(MSY) lead to a severe reduction of the total yield

2. If M is high it is difficult to estimate F(MSY) by the Y/R curve

This is not surprising: if M is high one must catch the fish before they die from natural causes, i.e. a high M requires a high F for optimum exploitation. If M is low it pays to let the fish grow to a large size before they are caught. When M is relatively low the gain in biomass from individual growth exceeds the loss caused by natural mortality for a long time of the cohort's life. When M is high the fish soon reach the age where losses due to natural mortality exceed the gain in biomass due to growth. These considerations make it reasonable to say that the stock is "growth overfished" if F exceeds F(MSY) (cf. Fig. 8.2.1).

In some cases (cf. Exercise 8.3) the Y/R-curve does not even have a maximum and an inexperienced management might come to the wrong conclusion that effort should be increased indefinitely. In such cases which are common in tropical fisheries, it is recommended not to look at the Y/R-curve, but at the biomass per recruit (B/R) curve, which is introduced in the next section. However, the two curves are not alternative and it is advisable always to plot them together.

The Y/R as a function of Tc, age at first capture, is closely related to the estimation of the optimum mesh size. We shall touch upon this subject in Exercises 8.3 and 8.4.

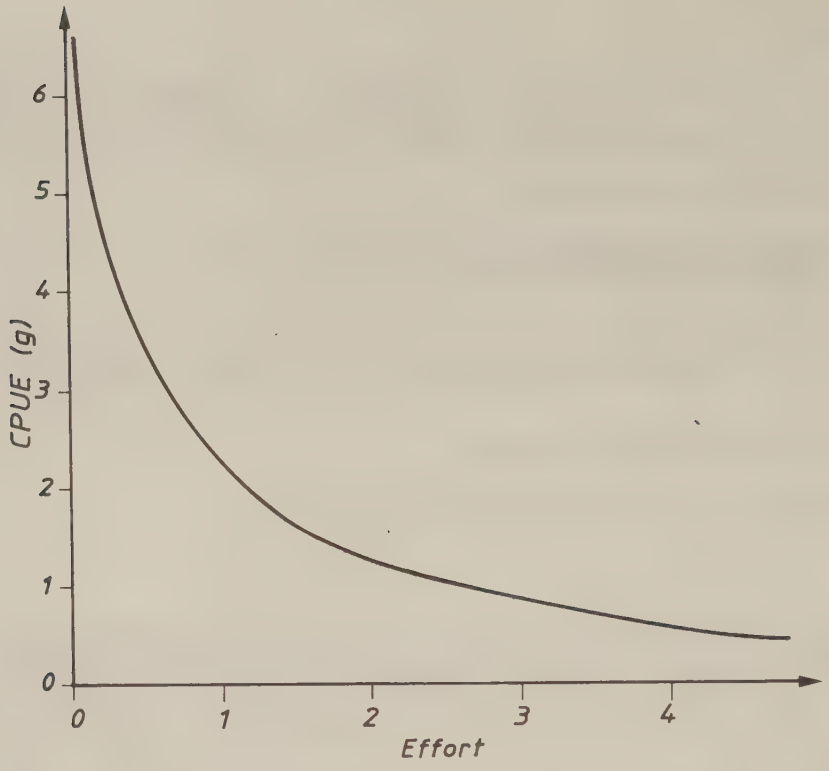


Fig. 8.3.1 Curve of CPUE (in weight) as a function of effort

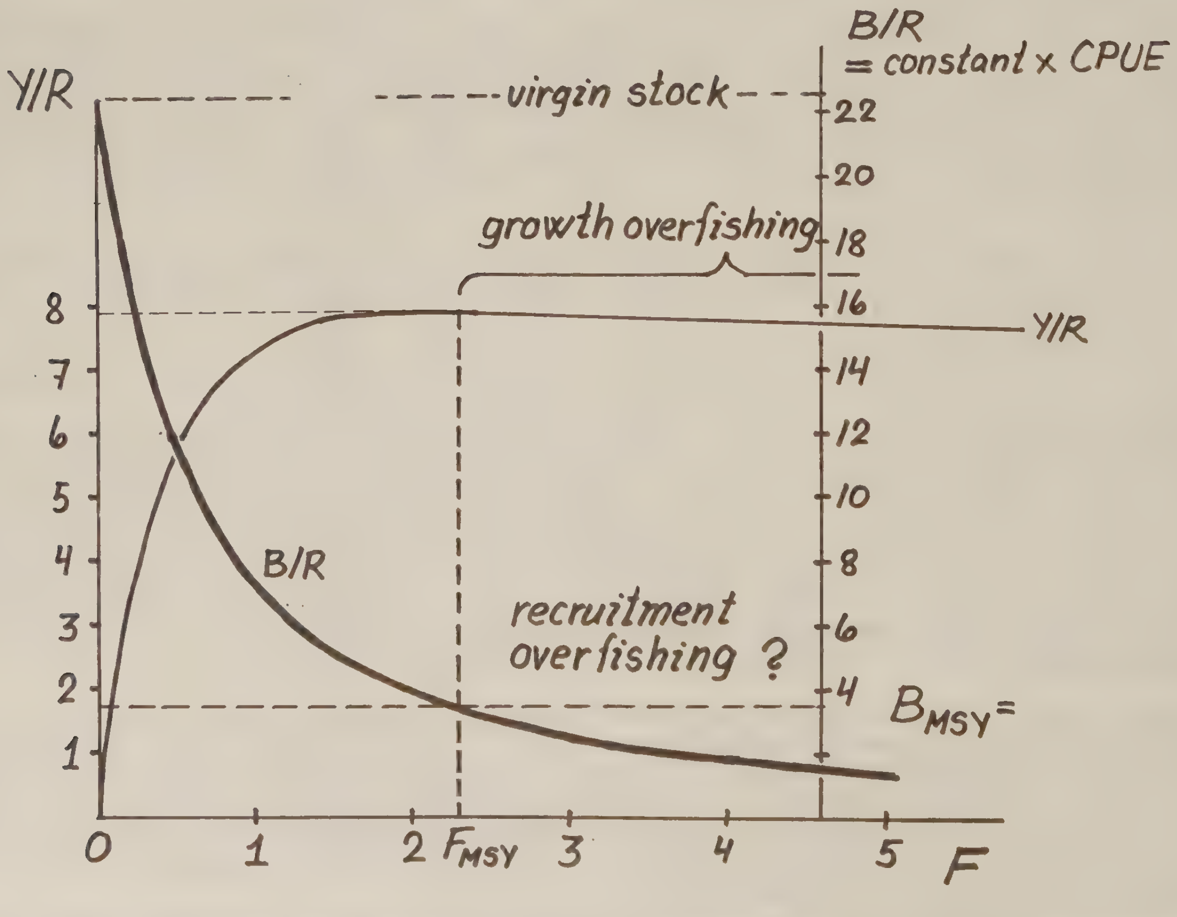


Fig. 8.3.2 Biomass per recruit curve corresponding to the Y/R-curve of Fig. 8.2.1 which is repeated here

8.3 Biomass per recruit

This model expresses the annual average biomass of survivors as a function of fishing mortality (or effort). The average biomass is related to the catch per unit of effort (cf. Section 4.3). Eq. 4.3.0.2 expresses the relationship between CPUE and numbers caught, which multiplied by the body weight gives:

$$\text{CPUE}(t)\ w(t) = q\ N(t)\ w(t)$$

or if $N(t)w(t)$ is replaced by $B(t)$, the symbol for biomass:

$$\text{CPUEw}(t) = q\ B(t) \qquad (8.3.1)$$

where CPUEw is the "weight of the catch per unit of effort". Thus, the biomass curve is expected to have the same shape as the CPUEw curve shown in Fig. 8.3.1.

The catch in numbers per year can be expressed as

$$C = F\ \bar{N}$$

(cf. Eq. 4.2.8 with t2-t1 = 1 year). By a similar argument it can be shown that the yield per year is

$$Y = F\ \bar{B}$$

where $\bar{B}$ is the average biomass in the sea during a year. It follows that:

$$\frac{\bar{B}}{R} = \frac{Y}{R}\frac{1}{F}$$

Because of the assumption of a constant parameter system the yield during one year is equal to the yield from a cohort during its lifespan.

Therefore we have the simple relationship between Y/R, (Eq. 8.2.4) and average biomass per recruit, $\bar{B}/R$:

$$\bar{B}/R = \frac{1}{F}\ Y/R \qquad (8.3.2)$$

Thus, it is easy to go from Y/R to $\bar{B}/R$.

If F = 0 there are problems. In that case we use:

$$\bar{B}/R = \exp[-M(Tc-Tr)]\ W_\infty \left[\frac{1}{Z} - \frac{3S}{Z+K} + \frac{3S^2}{Z+2K} - \frac{S^3}{Z+3K} \right] \qquad (8.3.3)$$

It is recommended always to calculate Y/R and $\bar{B}/R$ together. The easiest way to do so is to start by calculating B/R by Eq. 8.3.3 and then use

$$Y/R = F\ (\bar{B}/R)$$

to calculate Y/R (this follows from Eq. 8.3.2). The biomass per recruit as defined by Eq. 8.3.2 or Eq. 8.3.3 is the biomass of the exploited part of the cohort, i.e. the biomass of fish of age Tc or older.

Fig. 8.3.2 shows the B/R-curve corresponding to the Y/R of Fig. 8.2.1. It shows that for the threadfin bream the biomass corresponding to the optimum F-level F(MSY) is only 15% of the virgin stock biomass, Bv, the biomass of the unexploited stock.

Now one could ask the question whether such a reduction in spawning stock biomass influences the productivity of the stock. Obviously, there must be a level of spawning stock biomass below which the recruitment is reduced because of the reduced number of mature fish. These thoughts lead to a second overfishing concept, namely "recruitment overfishing". We say that a stock is "recruitment overfished" if the parent stock biomass is reduced to a level where there is a relationship between its biomass and the recruitment it can produce.

To illustrate this, consider Fig. 8.3.3. It shows the relationship between recruitment and the spawning stock biomass. As indicated by the questionmarks this relationship is not well understood. The only point known for sure is (0,0), i.e. with no parents there can be no offspring. As the parent stock starts to increase in size from (close to) zero the recruitment also increases. In the range of spawning stock biomass where there is such a stock/recruitment relationship with a positive slope we say that the stock is recruitment overfished.

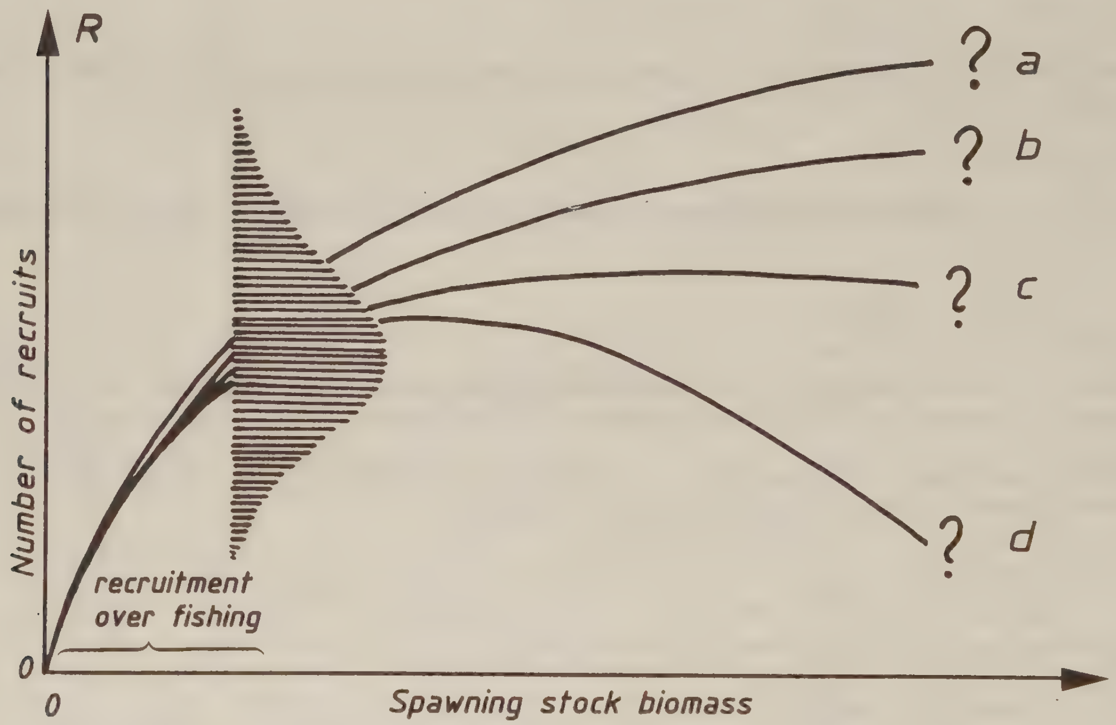


Fig. 8.3.3 **The stock/recruitment relationship in connection with the concept of "recruitment overfishing"**

However, this concept may not have much practical use, because the relationship between stock and recruitment is poorly understood. Very large variations of recruitment for given spawning stock size have been estimated for different stocks. Today we must conclude that the fisheries science is far from a solution to the stock/recruitment problem. More about this subject follows in Chapter 12.

Another, and sometimes more appropriate, way of using the B/R-curve is to interpret it as a CPUEw-curve. When managing a fishery, considerations on the possible income per boat are essential and this quantity of course is closely related to CPUEw. This (very important) aspect deals with fisheries economics, which, however, is not dealt with in this manual.
(See **Exercise(s)** in Chapter 17).

8.4 Beverton and Holt's relative Y/R concept

Beverton and Holt (1966) developed a more general version of their Y/R-model. The absolute value of Y/R expressed in grammes per recruit has no direct relation to fisheries management. What matters are the relative differences of Y/R for different values of F. An example of the kind of conclusion we can draw from the Y/R curve is: if F increases by 20% then yield reduces by 15%. As a consequence of this Beverton and Holt developed a version of the Y/R model in which the absolute Y/R-values have no meaning, but which has the advantage of requiring fewer parameters. This version is especially suitable for assessing the effect of mesh size regulations.

Let Lc be the 50% retention length, (L50%) i.e. the length at which 50% percent of the fish are retained by the gear if they enter it (cf. Fig. 6.1.0.1), and define the auxiliary variables:

$$U = 1 - Lc/L_\infty \qquad m = \frac{1 - E}{M/K} = \frac{K}{Z}$$

where E = F/Z is the exploitation rate or the fraction of deaths caused by fishing (cf. Section 4.2), M the natural mortality and K the von Bertalanffy parameter. U is the fraction of the growth to be completed after entry into the exploited phase.

The expression, (Y/R)', is defined by:

$$(Y/R)' = E\,U^m \left[1 - \frac{3U}{1 + m} + \frac{3U^2}{1 + 2m} - \frac{U^3}{1 + 3m} \right] \tag{8.4.1}$$

Eq. 8.4.1 is called the "relative Beverton and Holt yield per recruit model".

(Y/R)' is considered a function of U and E and the only parameter is M/K. The equation gives a quantity which is proportional to Y/R as defined by Eq. 8.2.4 as can be shown by a number of algebraic manipulations. It can be shown that $(Y/R)' = (Y/R)\exp[M(Tr-t_0)]/W_\infty$.

Fig. 8.4.1 shows the (Y/R)'-curve corresponding to the Y/R-curve of *Nemipterus marginatus*, Fig. 8.2.1, in the case of:

$$Lc = L(Tc) = L(1.0) = 28.4\,[1 - \exp(-0.37 * (1 + 0.2))] = 10.2 \text{ cm}$$

where L_∞ = 28.4 cm (see Section 3.3.6 and Fig. 3.3.6.1).

$$U = 1 - Lc/L_\infty = 1 - 10.2/28.4 = 0.641$$

As an example we calculate (Y/R)' for E = 0.5:

$$m = \frac{1 - 0.5}{1.1/0.37} = 0.168$$

$$(Y/R)' = 0.5 * 0.641^{0.168} * \left[1 - \frac{3 * 0.641}{1 + 0.168} + \frac{3 * 0.641^2}{1 + 2*0.168} - \frac{0.641^3}{1 + 3*0.168} \right]$$

$$= 0.047$$

(See **Exercise(s)** in Chapter 17).

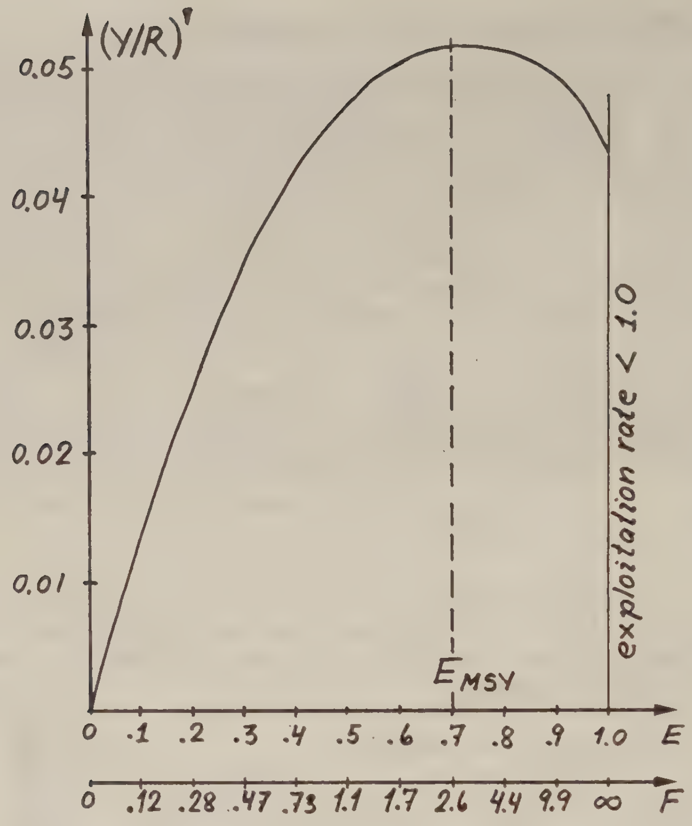


Fig. 8.4.1 Beverton and Holt's relative yield per recruit (Y/R)' curve corresponding to the Y/R-curve of Fig. 8.2.1 (Lc = 10.2 cm)

8.5 Yield per recruit from length data

Almost the same algebra as referred to in Section 8.4 transforms the equation for Y/R (Eq. 8.2.4) into a length based model (Marten, 1978). The original parameters and variables are F, M, W_∞, K, t_o, Tr and Tc. In the length transformed model we have L_∞, Lr and Lc in stead of t_o, Tr and Tc. The new equation is

$$Y/R = FA\ W_\infty \left[\frac{1}{Z} - \frac{3U}{Z+K} + \frac{3U^2}{Z+2K} - \frac{U^3}{Z+3K} \right] \qquad (8.5.1)$$

where U = 1 - Lc/L_∞ as in Eq. 8.4.1 and

$$A = \left[\frac{L_\infty - Lc}{L_\infty - Lr} \right]^{M/K}$$

Recalling that several methods of parameter estimation described in the preceding chapters give Z/K or M/K it may be of interest also to formulate Eq. 8.5.1 in such terms. Division by K outside the bracket, multiplication by K inside, and substituting z for Z/K gives:

$$Y/R = \frac{F}{K}\ A\ W_\infty \left[\frac{1}{z} - \frac{3U}{z+1} + \frac{3U^2}{z+2} - \frac{U^3}{z+3} \right] \qquad (8.5.2)$$

This equation contains F/K, M/K (in A) and Z/K (in z), has no reference to age and does not require a separate estimate of K.

9 ESTIMATION OF MAXIMUM SUSTAINABLE YIELD USING SURPLUS PRODUCTION MODELS

The objective of the models dealt with in this chapter is to estimate the optimum level of effort, i.e. the effort which produces the maximum sustainable yield (MSY). That was also the main objective of the Beverton and Holt model discussed in Chapter 8. However, the models of this chapter, the "surplus production models", are based on a different reasoning. The theory of surplus production models is reviewed in for example Ricker (1975), Caddy (1980), Gulland (1983) and Pauly (1984).

Beverton and Holt based their models on single fish considerations. Mortality and growth are separately modelled. In the surplus production models the stock is considered a big "lump" of biomass and no attempt is made to model on an age or length base. The surplus production models cannot deal with concepts like Tc ("age at first capture", cf. Section 8.1) and the effect of changes in mesh size, because these are related to the growth of individual fish. The death process which in all the foregoing sections has been modeled as a decay of the number of survivors is here considered a "negative growth" of the stock biomass. The cohort concept is not used at all.

The maximum sustainable yield is estimated from the following input data:

$f(i)$ = effort in year, $i = 1,2,\ldots,n$

(Y/f) = yield (catch in weight) per unit of effort in year i,

(Y/f) may be derived from the yield, $Y(i)$ of year i for the entire fishery and the corresponding effort, $f(i)$, by

$$(Y/f) = Y(i)/f(i), \quad i = 1,2,\ldots,n \qquad (9.1.1)$$

but they may also be obtained by direct observations of yield per unit of effort, on the basis of samples from the fishery.

Fig. 9.0.1 shows an example of a plot of $Y(i)/f(i)$ against $f(i)$. Data are from the demersal fishery off the North coast of Java during the years 1969-1977 (from Dwiponggo, 1979). In this case yield refers to the annual catch measured in units of 1000 tonnes and (Y/f) refers to yield per thousand standard vessels each year. Effort is given in units of standard vessels per year. Because the fleet consisted of a number of different boat types, the effort of each boat category has been converted to a standard unit before summing to obtain total effort. Not surprisingly, the trend of (Y/f) shows a decline for increasing effort. Considering the stock biomass as a limited resource shared by the boats participating in the fishery, we expect a smaller share per boat the more boats there are.

9.1 The Schaefer model and the Fox model

The simplest way of expressing yield per unit of effort as a function of the effort is the linear model suggested by Schaefer (1954):

$$\text{Schaefer:} \quad Y(i)/f(i) = a + b\,f(i) \quad \text{if } f(i) \le -a/b \qquad (9.1.2)$$

Eq. 9.1.2 is called the "Schaefer model".

The slope, b, must be negative if (Y/f) decreases for increasing f (see Fig. 9.1.1). The intercept, a, is the (Y/f) value obtained just after the first boat fishes on the stock for the first time. The intercept therefore

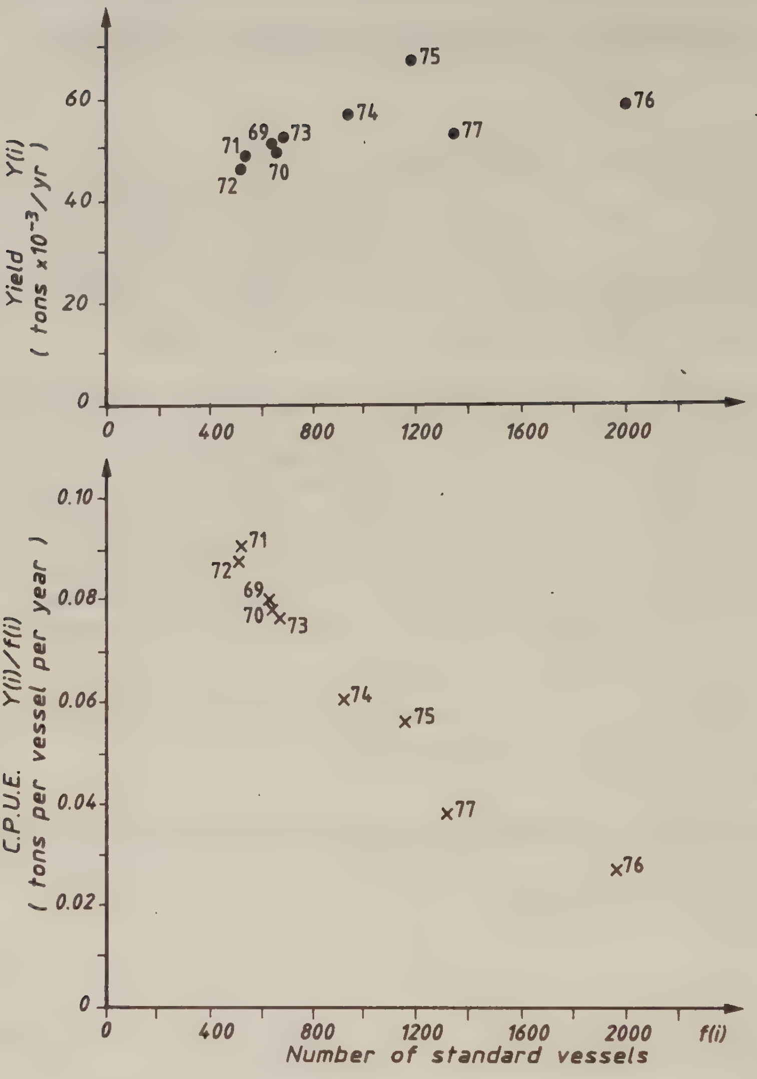


Fig. 9.0.1 Trends of yield, Y(i), and yield per unit of effort, Y(i)/f(i), off the North coast of Java (based on data from Table 9.1.1)

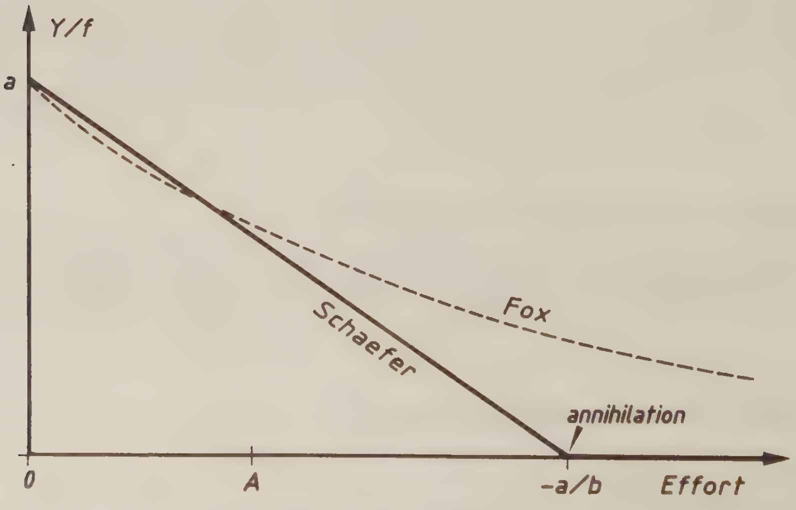


Fig. 9.1.1 Illustration of the different assumptions behind the Schaefer model and the Fox model

must be positive. Thus, $-a/b$ is positive and (Y/f) is zero for $f = -a/b$. As a negative value of (Y/f) is absurd the model only applies to f-values less than $-a/b$.

An alternative model was introduced by Fox (1970). It gives a straight line when the logarithms of (Y/f) are plotted on effort:

$$\ln (Y(i)/f(i)) = c + d\ f(i) \tag{9.1.3}$$

Eq. 9.1.3 is called the "Fox model".

Both models conform to the assumption that (Y/f) declines as effort increases, but they differ in the sense that the Schaefer model implies an effort level for which (Y/f) equals zero, namely $f = -a/b$ (cf. Fig. 9.1.1) whereas in the Fox model:

$$Y(i)/f(i) = \exp(c + d\ f(i)) \tag{9.1.4}$$

(Y/f) is greater than zero for all values of f (cf. Fig. 9.1.1).

Bearing in mind that (Y/f) is proportional to the biomass (cf. Section 8.2) we can write

$$(Y/f) = q\ B = \begin{cases} a + b\ f(i) & \text{(Schaefer)} \\ \exp(c + d\ f(i)) & \text{(Fox)} \end{cases}$$

Fig. 9.1.1 illustrates the basic features of the two models. For f close to zero (Y/f) takes the maximum value and so does the biomass. The biomass corresponding to $f = 0$ is called the "virgin stock biomass" or the "unexploited biomass", denoted by Bv. Thus, replacing Y/f by qBv in Eqs. 9.1.2 and 9.1.4 gives:

$$q\ Bv = \begin{cases} a & \text{(Schaefer)} \\ \exp(c) & \text{(Fox)} \end{cases} \quad \text{or} \quad Bv = \begin{cases} a/q & \text{(Schaefer)} \\ \exp(c)/q & \text{(Fox)} \end{cases}$$

(see Fig. 9.1.1). The Bv for the two models must be the same. When increasing f from zero to level A (see Fig. 9.1.1) the two curves are approximately equal, but to the right of A the differences become larger. Thus, the choice between the two models becomes important only when relatively large values of f are observed.

It cannot be proved that one of the two models is superior to the other. You may choose the one you believe is the most reasonable one in each particular case or the one which gives the best fit to the data. The Beverton and Holt model is more in agreement with the Fox model, because both have a curvilinear relationship between catch per unit effort (or biomass per recruit) and effort (or F). Compare Figs. 8.3.1 and 9.1.1.

The estimation procedure for the parameters (Schaefer: a and b, Fox: c and d) is set out in Table 9.1.1 for the example of Fig. 9.0.1. For the regression analysis, recall Section 2.4 and Exercise 2.4.

The next step is to derive an estimate of MSY from the estimated parameters of both models. To that purpose we rewrite Eqs. 9.1.2 and 9.1.4 expressing the yield as a function of effort:

$$\text{Schaefer:} \begin{cases} Y(i) = a\ f(i) + b\ f(i)^2 & \text{if } f(i) < -\ a/b \\ Y(i) = 0 & \text{if } f(i) \geq -\ a/b \end{cases} \tag{9.1.5}$$

$$\text{Fox:} \quad Y(i) = f(i) \exp[c + d\ f(i)] \tag{9.1.6}$$

Table 9.1.1 **The calculation procedure for estimating the MSY and f(MSY) by the Schaefer model and by the Fox model using catch and effort data from the trawl fishery off the North coast of Java (Dwiponggo, 1979)**

year i	yield (000'tons) Y(i)	effort (nos. of standard vessels) f(i) x	SCHAEFER Y/f y	FOX ln(Y/f) y
1969	50	623	0.080	-2.523
1970	49	628	0.078	-2.551
1971	47.5	520	0.091	-2.393
1972	45	513	0.088	-2.434
1973	51	661	0.077	-2.562
1974	56	919	0.061	-2.798
1975	66	1158	0.057	-2.865
1976	58	1970	0.029	-3.525
1977	52	1317	0.039	-3.232
mean value	:	923.22	0.0667	-2.7648
standard deviation	:	485.14	0.02171	0.3873

intercept (Schaefer: a, Fox: c) :	0.1065	-2.0403
slope (Schaefer: b, Fox: d) :	-0.00004312	-0.0007848
variance of slope: $sb^2 = ((sy/sx)^2 - b^2)/(9-2)$ *) :	$2.041*10^{-11}$	$3.000*10^{-6}$
standard deviation of slope, sb *) :	0.00000452	0.0000556
confidence limits of slope: b + t(9-2) sb (t(7) = 2.37) :	-0.000032 -0.000053	-0.00065 -0.00092
variance of intercept: $sa^2 = sb^2 (sx^2 (n-1)/n + x^2)$ *) :	0.00002167	0.003277
standard deviation of intercept :	0.0049	0.0572
confidence limits of intercept : a + t(9-2) sa *) :	0.095 0.118	-2.18 -1.90
MSY: Schaefer: $- a^2/(4b)$:	65.8	
Fox : $-(1/d) \exp(c-1)$:		60.9
f(MSY): Schaefer: $- a/(2b)$:	1235	
Fox : $- 1/d$:		1274

*) a,b replaced by c,d for the Fox-model

The graph of Eq. 9.1.5, the Schaefer model (see Fig. 9.1.2 and Table 9.1.1) is a parabola and it takes its maximum value for:

$$f(MSY) = - 0.5\ a/b \tag{9.1.7}$$

and the corresponding maximum sustainable yield becomes:

$$MSY = - 0.25\ a^2/b \tag{9.1.8}$$

For Eq. 9.1.6, the Fox model (see Fig. 9.1.2 and Table 9.1.1) we get:

$$f(MSY) = - 1/d \tag{9.1.9}$$

$$MSY = - (1/d) \exp(c - 1) \tag{9.1.10}$$

(Eqs. 9.1.7 to 9.1.10 are derived by differentiating Y with respect to f, and solve dY/df = 0 for f).

(See **Exercise(s)** in Chapter 17).

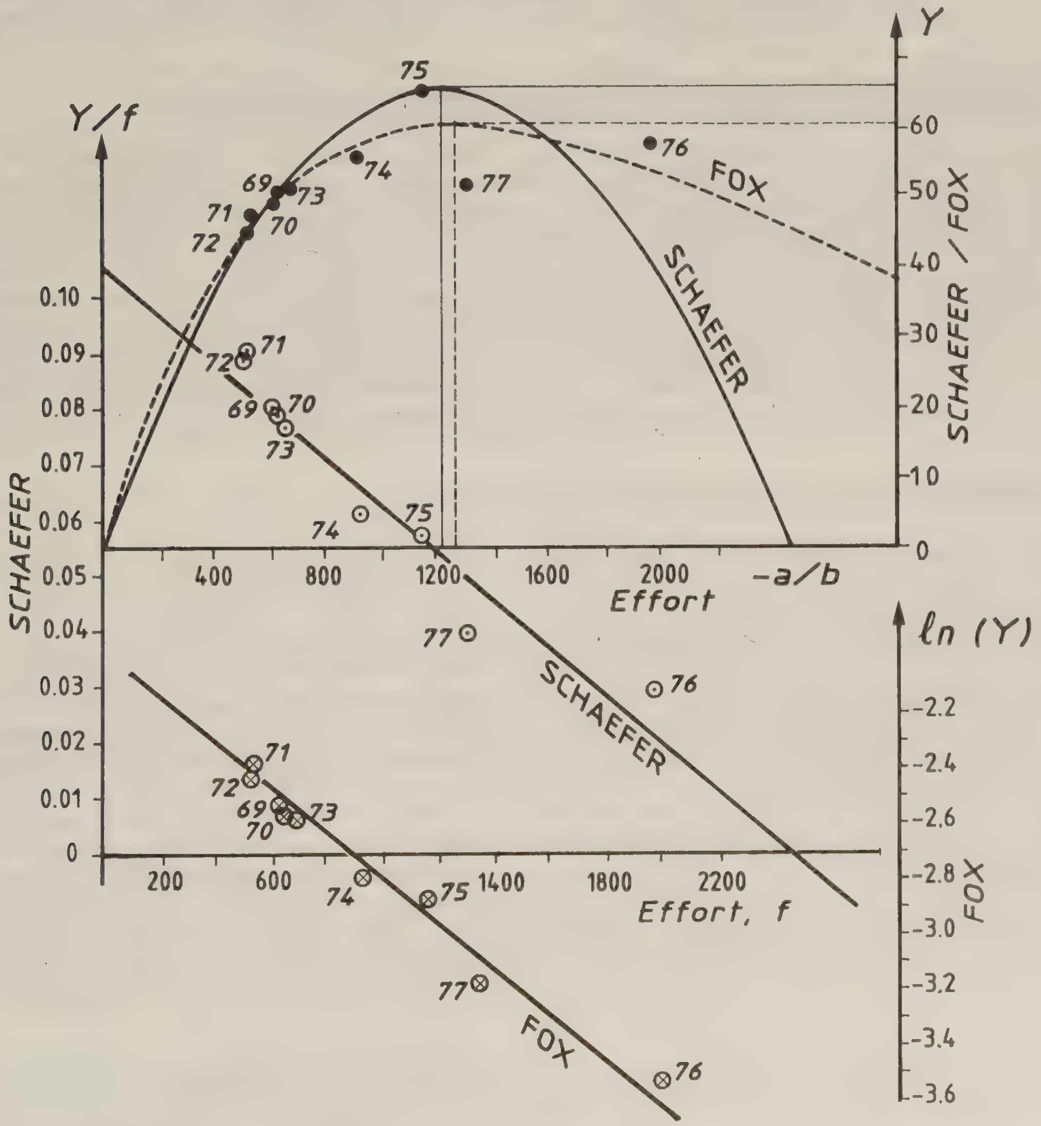


Fig. 9.1.2 The Schaefer and the Fox models illustrated by the demersal fisheries off the North coast of Java (based on data from Table 9.1.1)

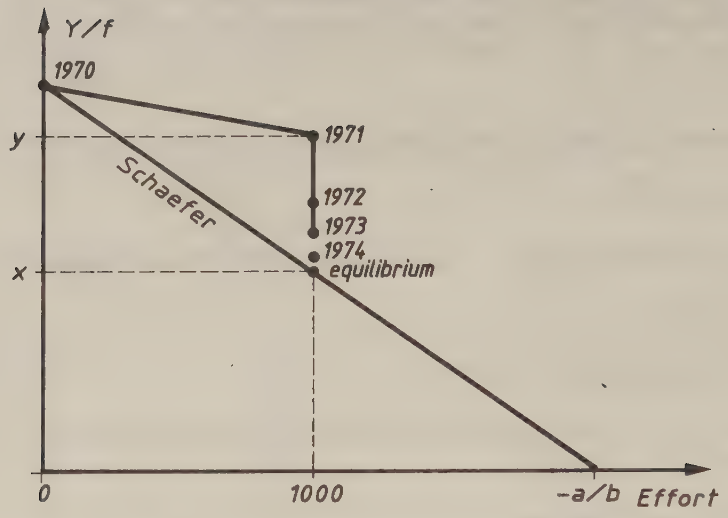


Fig. 9.2.1 Illustration of the concepts of an equilibrium situation and the transition period (for further explanations, see text)

9.2 The assumptions behind the surplus production models

So far we have dealt mainly with the computational procedure, which was straight forward. One could now ask why we should be concerned with complicated models like the Beverton and Holt yield per recruit model when estimates of MSY can be obtained so easily from the surplus production models. One answer is that what we gain in simplicity with the surplus production models has the cost of a number of assumptions about the dynamics of fish stocks, which may be (and nearly always are) impossible to justify. Assumptions which we know are never strictly fulfilled, but which we hope are approximately met. Some of these assumptions are discussed below. The reasoning is based on the Schaefer model, but applies to the Fox model too.

THE ASSUMPTION OF AN EQUILIBRIUM SITUATION:

To explain this concept we consider a situation in which a virgin stock starts to be exploited in, say, 1971 (see Fig. 9.2.1), by, say, 1000 boats. Suppose the Schaefer line in Fig. 9.2.1 applies to this stock. According to the model the yield in 1971 should be x (see Fig. 9.2.1).

However, it turned out to be y, i.e. a larger value than predicted by the model. This is because when fishing started in 1971 the biomass was still Bv (the virgin stock biomass) and only after a certain period of exploitation the biomass declines.

When fishing continued in 1972 the biomass was reduced due to the removal by fishing in 1971 and the 1972 catch therefore became smaller than that of 1971. Each year the resource is reduced, the reduction being smaller the longer time has elapsed since the introduction of the 1000 boats. Eventually, the system will stabilize at the Y/F-level x. We say that the system has reached a new "equilibrium situation" after a "transition period".

For the equilibrium situation the production of biomass per time unit, $\Delta B/\Delta t$, equals the removal by fishing, the yield per time unit, $\Delta Y/\Delta t$:

$$\frac{\Delta B}{\Delta t} = \frac{\Delta Y}{\Delta t} \qquad (9.2.1)$$

The "equilibrium situation" in the surplus production models is comparable to the "stabilized constant parameter system" in the Beverton and Holt models (cf. Section 8.1).

THE BIOLOGICAL ASSUMPTIONS:

The biological reasoning behind the model was adequately formulated by Ricker (1975):

> "1. Near maximum stock density, efficiency of reproduction is reduced, and often the actual number of recruits is less than at smaller densities. In the latter event, reducing the stock will increase recruitment.
>
> 2. When food supply is limited, food is less efficiently converted into fish flesh by a large stock than by a smaller one. Each fish of the larger stock gets less food individually; hence a larger fraction is used merely to maintain life, and a smaller fraction for growth.
>
> 3. An unfished stock tends to contain more older individuals, relatively, than a fished stock. This makes for decreased production, in at least two ways:
>
> (a) Larger fish tend to eat larger foods, so an extra step may be inserted in the food pyramid, with consequent loss of efficiency of utilization of the basic food production.
>
> (b) Older fish convert a smaller fraction of the food they eat into new flesh - partly, at least because mature fish annually divert much substance to maturing eggs and milt".

However, it is also possible to consider these models as purely empirical. For instance, if observations of (Y/f) plotted on f gives a curve complying with the Fox model, this model may be applied without any concern for a possible biological explanation.

ASSUMPTIONS ON THE CATCHABILITY COEFFICIENT:

We assume that fishing mortality is proportional to effort (cf. Eq. 4.6.1):

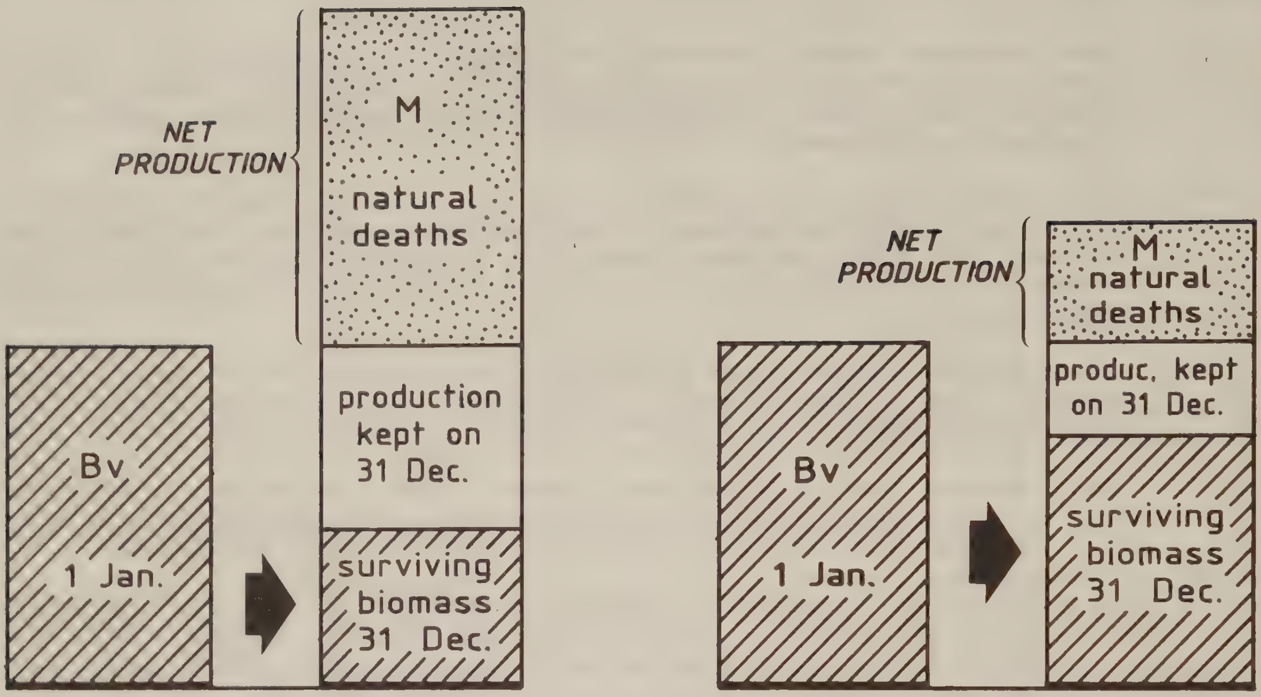
$$F = q\ f$$

This assumption in itself is not controversial (if f is a reasonable measure of effort). The problems come when f is measured in, for example the number of boat days per year over a series of years. In most cases the efficiency of the boats has changed over a long period; often the boats have become larger and better equipped. Thus, 100 boat days in, say, 1978 may create a larger fishing mortality than 100 boat days in 1968. This means that q becomes a function of time or rather, a function of the technical development which is usually a function of time. It has proved very difficult to account for changes in q, caused by increased fishing efficiency and usually it is assumed that q remains constant. Therefore, one should be cautious not to include too long a time series of data in the surplus production analysis. Alternatively, changes of q must be taken into account.

9.3 Gulland's empirical formula

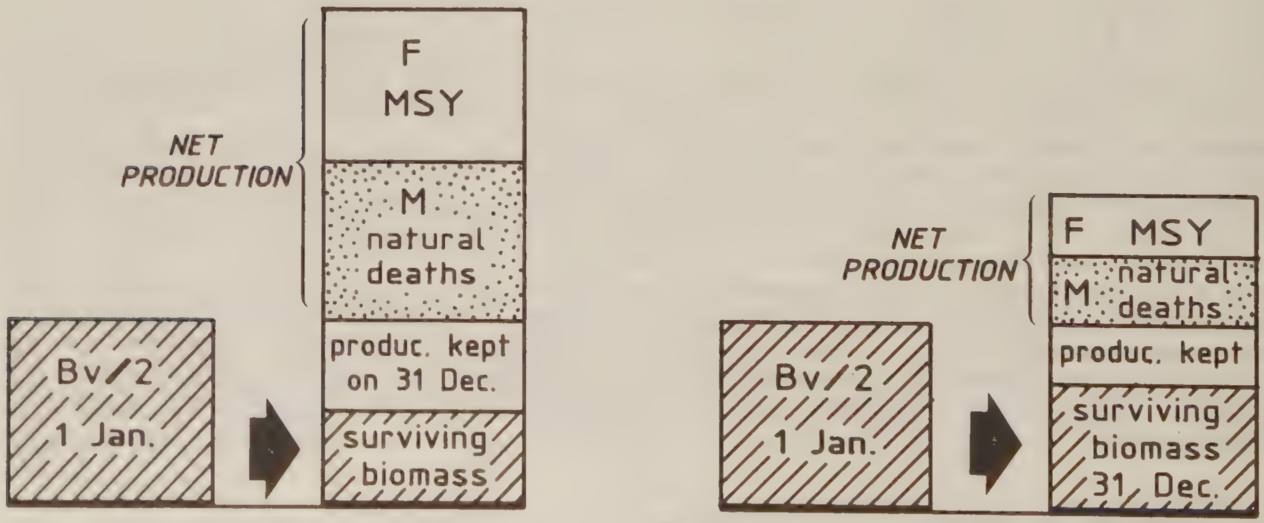
In this section we consider the case of poorly investigated stocks. Time series of catch and effort data are not available, but some estimates of overall biomass and the natural mortality are present.

EQUILIBRIUM CONDITIONS


NET PRODUCTION
M
natural deaths
production kept on 31 Dec.
Bv
1 Jan.
surviving biomass 31 Dec.
NET PRODUCTION
M
natural deaths
produc. kept on 31 Dec.
Bv
1 Jan.
surviving biomass 31 Dec.

A1: Virgin stock (Bv)
High M
High net production

B1: Virgin stock (Bv)
Low M
Low net production


NET PRODUCTION
F
MSY
M
natural deaths
produc. kept on 31 Dec.
Bv/2
1 Jan.
surviving biomass
NET PRODUCTION
F MSY
M natural deaths
produc. kept
Bv/2
1 Jan.
surviving biomass 31 Dec.

A2: Exploited stock (Bv/2)
High M
High net production

B2: Exploited stock (Bv/2)
Low M
Low net production

Fig. 9.3.1 Illustration of Gulland's formula MSY = 0.5 M Bv (for further explanation, see text)

Let Bv be the virgin stock biomass and let M be the natural mortality. Gulland (1971) suggested the following way of estimating maximum sustainable yield:

$$MSY = 0.5\ M\ Bv \qquad (9.3.1)$$

This formula has been used especially on sparsely investigated and lightly exploited stocks. Bv is often estimated by the swept area method (Chapter 13) and M is often a value estimated for a sea area which is believed to be similar to the one under investigation. As Gulland's formula requires an estimate of the virgin stock biomass, Bv, it is in practice applicable only to unexploited stocks. There is no proper scientific justificaiton for Eq. 9.3.1 (Gulland, pers. comm.). However, the following statements which were made already by Tiurin (1962) and Alverson and Pereyra (1969) appear reasonable:

1. MSY must depend on the virgin stock biomass, Bv
2. A high M corresponds to a high production (this is further discussed below)
3. If the biomass = 0.5 Bv and F = M under optimum exploitation Eq. 9.3.1 is fulfilled.

Item 2 above is illustrated in Fig. 9.3.1. We consider the annual biomass budget for two virgin stocks, A.1 with a high production rate and B.1 with a low production rate. For stock A.1 the loss caused by natural mortality is large and to maintain the stock the net production must be of equal size. For stock B.1 the loss due to natural deaths is small and only a small net production is required to counter-balance this loss. The cases A.2 and B.2 deal with the same two stocks but after they have been for a while under optimum exploitation i.e., produce the maximum sustainable yield. In A.2 a large part of the natural deaths has been replaced by deaths due to fishing. This is also the case for B.2, but the potentials for replacing natural deaths by deaths due to fishing is smaller in case B.2 than in case A.2.

Although widely used, Eq. 9.3.1 has been criticized by a number of workers. Caddy and Csirke (1983) showed that the third assumption (F = M under optimum exploitation) does not apply to many stocks, especially prey species (e.g. shrimps). Beddington and Cooke (1983) concluded, based on simulation studies that Eq. 9.3.1 generally overestimates MSY by a factor 2 to 3. Thus, replacing "0.5" by "0.2" in Eq. 9.3.1 might perhaps give a better (and consequently much lower) estimate of MSY.

9.4 Cadima's formula

A generalized version of Gulland's estimator was proposed by Cadima (in Troadec, 1977) for exploited fish stocks for which only limited stock assessment data are available.

Cadima's estimator has the form:

$$MSY = 0.5\ Z\ B \qquad (9.4.1)$$

where B is the average (annual) biomass and Z the total mortality. Since Z = F + M and Y = F B, the author suggested that in the absence of data on Z Eq. 9.3.2 could be rewritten:

$$MSY = 0.5\ (Y + M\ B) \qquad (9.4.2)$$

where Y is the total catch in a year and B is the average biomass in the same year.

As most stocks are now already exploited this equation is more frequently used in developing and some developed fisheries where catch and effort time series are not yet available, but biomass estimates are occasionally obtained from, for instance, trawl or acoustic surveys.

9.5 MSY estimates based on the surplus production model

Based on the considerations in the foregoing section, Garcia, Sparre and Csirke (1989) suggested two alternative ways to estimate the potential yield of exploited fish stocks which have basically the same foundation and applications as the Gulland and Cadima estimators, but which are consistent with the underlying models. The two proposals have been derived based on the Schaefer model and the Fox model.

Both methods assume that the observations:

$\bar{B}$, average biomass and Y (current yield)

are available for one year only. They also assume that natural mortality, M, is known and they assume a relationship between M and F(MSY) of the form:

$$M = k\,F(MSY) \qquad (9.5.1)$$

where k is a constant.

As $F = Y/\bar{B}$ and $Y/F = \bar{B}$, we can write the surplus production models in the form:

$$\text{Schaefer:} \quad \bar{B} = a - b\,(Y/\bar{B}) \qquad (9.5.2)$$

$$\text{Fox:} \quad \ln \bar{B} = c - d\,(Y/\bar{B}) \qquad (9.5.3)$$

Suppose observations B1 and Y1 are available and we have a "guesstimate" of M, then combined with the assumption of Eq. 9.5.1 we get:

$$\text{Schaefer:} \quad B1 = a - b\,(Y1/B1) \quad (f(MSY) = k\,M = -\,a/2b) \qquad (9.5.4)$$

$$\text{Fox:} \quad \ln B1 = c - d\,(Y1/B1) \quad (f(MSY) = k\,M = -\,1/d) \qquad (9.5.5)$$

These equations can be solved for a and b, c and d in the Schaefer and the Fox model, respectively:

$$\text{Schaefer:} \quad a = \frac{2\,F(MSY)\,B1^2}{2\,F(MSY)\,B1 - Y1} \quad ; \quad b = \frac{B1^2}{2\,F(MSY)\,B1 - Y1} \qquad (9.5.6)$$

$$\text{Fox:} \quad c = \ln B1 + Y1/(B1\,F(MSY)) \quad ; \quad d = -1/F(MSY) \qquad (9.5.7)$$

Once we have (a,b) or (c,d) we can estimate the MSY by Eq. 9.1.8 and 9.1.10:

$$\text{Schaefer:} \quad MSY = -\,0.25\,a^2/b$$

$$\text{Fox:} \quad MSY = -\,(1/d)\,\exp(c-1)$$

and draw the yield curves (cf. Eqs. 9.1.5 and 9.1.6). The MSY's in terms of B, Y and F(MSY) are given below.

Thus, one pair of observations (B1,Y1) and assumptions on M and the relationship between M and F(MSY) (F(MSY) = kM) are sufficient information to get a first rough estimate of the yield curve (Schaefer: a,b or Fox: c,d) from which a first rough estimate of MSY can be derived.

The MSY expression corresponding to the Schaefer model is found by inserting Eq. 9.5.6 into Eq. 9.1.5:

$$\mathrm{MSY} = \frac{F(\mathrm{MSY})^2\ \bar{B}^{\ 2}}{2\ F(\mathrm{MSY})\ \bar{B} - Y} \qquad (9.5.8)$$

The yield curve is determined by a and b (Eq. 9.5.6). When F(MSY) is not known (as is most often the case) it may be replaced by M or kM. In the case F(MSY) = M we get:

$$\mathrm{MSY} = \frac{M^2\ \bar{B}^{\ 2}}{2\ M\ \bar{B} - Y} \qquad (9.5.9)$$

If the stock is unfished (i.e. when F = 0, Y = 0 and B = Bv) Eq. 9.5.9 becomes Gulland's original formula (Eq. 9.3.1).

If the stock in question responds better to the Fox production model (Eq. 9.1.6), we get the expression:

$$\mathrm{MSY} = F(\mathrm{MSY})\ \bar{B}\ \exp\ (\ Y/(F(\mathrm{MSY})\ \bar{B}) - 1) \qquad (9.5.10)$$

The yield curve is determined by c and d (Eq. 9.5.7). In the special case where F(MSY) = k M with k = 1 (or F(MSY) = M) Eq. 9.5.10 becomes:

$$\mathrm{MSY} = M\ \bar{B}\ \exp\ (\ Y/(M\ \bar{B}) - 1) \qquad (9.5.11)$$

The interested reader is referred to the original paper (Garcia, Sparre and Csirke, 1989) for further details.

9.6 Validation of MSY-estimates based on empirical formulas

When working with mathematical models the fishery biologist should check whether the basic assumptions of the models are fulfilled. This applies in particular to the above mentioned empirical formulas (Garcia, Sparre and Csirke, 1987 and 1989). For example, the biomass (B) is meant to be the exploited average biomass, and both the catch and the biomass referred to should have the same age (or size) structure. For instance, the biomass figure should not include small sizes which are not available to the fishery. This biomass is the annual biomass value and seasonal oscillations caused by change in growth, mortality or recruitment, which are likely to be more important in short-lived species such as shrimps, squids and anchovies, should be taken into account and as far as possible be leveled off to obtain an appropriate annual average of the total biomass.

If seasonal oscillations of biomass are caused by migration, then the peak biomass representing the real size should be used. If another country is exploiting the same stock at another season of the year, in another area, the combined catch should be included in the calculations and the MSY estimate should refer to the whole unit stock. The estimated potential yield could also be validated by comparison with other stocks for which better information might be available.

Some of the questions to be asked are: How does the estimate of MSY/km² stand with respect to the estimates for other stocks of the same species in ecological similar areas exploited under similar fishing regimes? Does the size structure of the catch provide support for assessment implying that the stock is heavily overfished (e.g. predominance of juveniles) or underfished (e.g., predominance of large, old fish)?

It should be noted that a closer look at the length frequency composition of the catch would give some guidance on relative levels of exploitation and this should be among the first data to be collected in any developing fishery in order to allow some estimation of the total mortality rate for cross-checking with other methods (cf. Chapters 3 and 4).

If the stock in question has been exploited for some time it is likely that a time series of catch data is available, which should also be examined. Even if no detailed effort data are available, the indication that after a period of sustained increase, the total catch has been stable for some time may mean that the MSY has been reached at least for the present regime of exploitation, while if the catch has dropped from a previous high level it may mean that the stock is overfished and an average of the highest catches experienced in the past may provide an independent approximation to the MSY. In interpreting catch time series as suggested above, one assumes that such variations in catches are caused by changes in fishing effort and not by causes of environmental or socio-economic origin.

9.7 Munro and Thompson's plot

The surplus production models are usually applied to time series of CPUE and effort. Munro and Thompson (1983 and 1983a), however, applied the surplus production model to a set of data from the Jamaica fishery all collected in the same year but representing different fishing grounds. Fig. 9.7.1 shows a schematic map of Jamaica divided into 9 areas. Except for area B they coincide with the parishes of Jamaica. The fishery considered in Munro and Thompson (1983) is the Jamaican coral reef fishery (mainly a local trap fishery operated from canoes). Coral reef fish are considered little mobile and it was assumed that each area (Fig. 9.7.1) has its own stocks which are independent of the neighbouring stocks (little mixing). Thus, the basic assumption is that the "ecological regimes" do not differ substantially around the island. Based on that assumption it makes sense to plot CPUE on the effort of one year by fishing area. For a stock distributed over all Jamaican waters such an approach should not be used.

Table 9.7.1 shows the CPUE and effort data for the Jamaican shelf fishery on shelf-dwelling species (excluding oceanic pelagics) in 1968.

Effort is expressed in units of canoes per km^2 per year to accommodate the assumption that each area has the same relative potentials, i.e. can support the same production per area unit. Thus, if exploited at the same rate (same effort per unit area per year) all areas should yield the same number of kg per unit area per year (cf. column D of Table 9.7.1).

The yield per unit area was based on the area of the shelf and the proximal banks of each parish. For further details the reader is referred to the original paper (Munro and Thompson, 1983 and 1983a).

The relationship between CPUE and effort is here assumed to be of the Fox type (Eq. 9.1.3). Fig. 9.7.2 shows the plot of ln(CPUE) on effort as well as the yield per unit area. Munro and Thompson excluded area F from the regression analysis, but did not explain why.

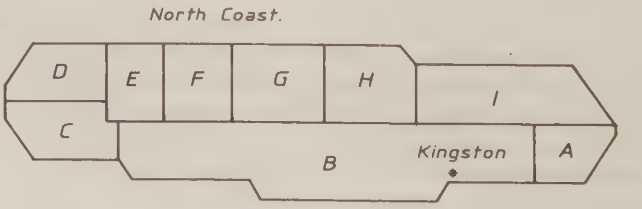


Fig. 9.7.1 Schematic map of Jamaica showing the parishes used for the "Munro and Thompson plot" (Fig. 9.7.2)

Table 9.7.1 **Input data for Munro and Thompson Plot (for further details, see text). From Munro and Thompson (1983a)**

	A	B	C = lnB	D = AB
parish (stock)	effort canoes per km²	CPUE kg per canoe per year	ln(CPUE) Fox model	yield kg per km² per year
A	1.63	2367	7.769	3858
B	0.38	3279	8.095	1246
C	3.09	1407	7.249	4348
D	5.63	556	6.321	3130
E	4.43	974	6.881	4315
(F)	5.51	1306	7.175	7196
G	4.58	564	6.335	2583
H	4.20	767	6.642	3221
I	1.49	1875	7.536	2794

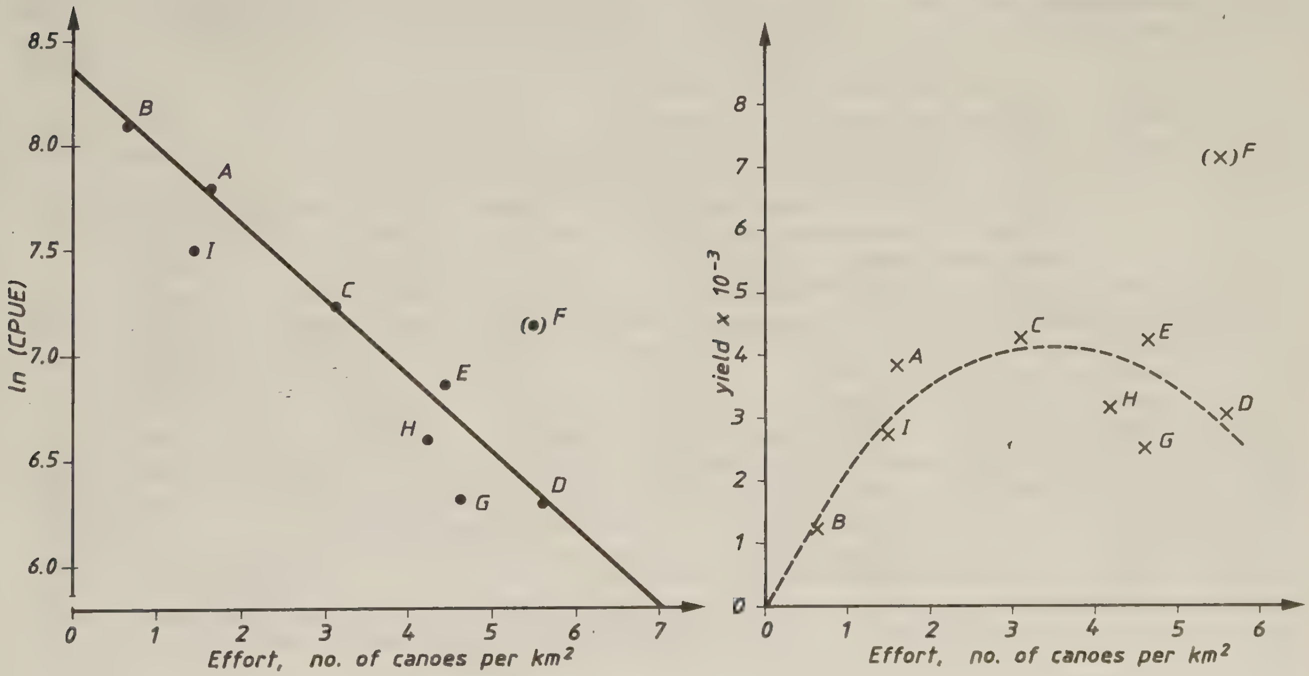


Fig. 9.7.2 Munro and Thompson plot. Based on data in Table 9.7.1 (from Munro and Thompson, 1983)

9.8 Standardization of effort

In Section 4.3 it was suggested that effort is proportional to fishing mortality. This is true, of course, if we define effort as something proportional to fishing mortality, but such a definition has no practical application. In practice we will have to choose a measure for effort we believe is related to fishing mortality or rather fishing power. There are many possible choices. For a trawl fishery we may consider:

Table 9.8.1 Example (hypothetical) to illustrate summation of effort for different effort units

year	PURSE SEINE			R1= rel.	BEACH SEINE			R2= rel.
y	Y1(y)	f1(y)	Y1/f1	Y/f	Y2(y)	f2(y)	Y2/f2	Y/f
1971	100	10	10.00	1.089	300	800	0.375	1.047
1972	200	21	9.52	1.037	250	669	0.374	1.043
1973	400	43	9.30	1.013	200	543	0.368	1.028
1974	700	81	8.64	0.941	150	430	0.349	0.974
1975	1200	142	8.45	0.920	100	307	0.326	0.909
Mean Y/f			9.18				0.358	

year	POLE AND LINE			R3= rel.	TROLLING			R4= rel.
y	Y3(y)	f3(y)	Y3/f3	Y/f	Y4(y)	f4(y)	Y4/f4	Y/f
1971	40	10000	0.00400	1.11	350	200000	0.00175	1.04
1972	80	20900	0.00383	1.06	344	200000	0.00172	1.02
1973	120	31100	0.00386	1.07	339	200000	0.00170	1.01
1974	80	23200	0.00345	0.96	333	200000	0.00167	0.99
1975	40	13700	0.00292	0.81	320	200000	0.00160	0.95
Mean Y/f			0.00361				0.00169	

Yi(y) = yield of gear i in year y. i = 1,..,4. y = 1970,...,1975.
fi(y) = effort of gear i in year y.
Yi(y)/fi(y) = catch per unit of gear i effort in year y.
Ri(y) = relative CPUE = CPUE(y)/(Mean CPUE)

year	total Y sampled YS(y)	weighted sum ΣYi * Ri	relative CPUE R(y)	total yield YT(y)	relative effort YT(y)/R(y)	normalized rel. effort E(y)
1971	790	830	1.052	7900	7515	0.653
1972	874	904	1.034	8740	8451	0.734
1973	1059	1080	1.020	10590	10386	0.903
1974	1263	1210	0.958	12630	13183	1.146
1975	1660	1531	0.9224	16600	17997	1.564
			Mean YT/R		11506	

$YS(y) = \sum_{i=1}^{4} Yi(y)$ = sum of yields of gears for which effort is known (yield of sampled gears)

$R(y) = \sum_{i=1}^{4} Yi(y)*Ri(y)/YS(y)$ = sum of relative CPUE weighted by the yields in year y

YT(y) = total yield of all gears (including gears for which effort is not known)

YT(y)/R(y) = relative effort of year y

$E(y) = \frac{YT(y)/R(y)}{\text{mean } YT/R}$ = normalized relative effort of year y

number of trawlers
number of trawler days
number of standard trawlers (taking into account the boat type)
number of standard trawler days
... etc.

For a handline fishery it may be more appropriate to consider the number of fisherman days or the number of hooks used times the number of days. In this case it may be necessary to take into account that the fishermen on the same boat compete, so that effort is not a linear function of the number of fishermen.

In general, a measure which can be shown to be linearly related to the catch rate is a suitable measure. That is, if it can be shown that two units of effort catch twice as much as one unit of effort when operating under equal conditions the effort measure is a suitable one. For example, the number of fishing hours times engine horse power may be a suitable measure of effort in some trawl fishery, whereas in a gill net fishery the boat type and the number of hours are likely to be less important than the number of gill nets set per day. In both of the above mentioned cases the number of fishermen may not be linearly related to the fishing power.

There are considerable difficulties in defining suitable measures of effort for a single gear as discussed above. When trying to define the effort for combinations of gears exploiting the same resources some rather intricate problems are encountered. The methods for standardization of effort units defined below assumes that all the units defined are suitable.

Before we start the discussion on standardization of effort units we notice that:

$$\frac{\text{Yield}}{\text{CPUE}} = \text{effort} \quad \text{as} \quad \text{CPUE} = \text{yield/effort}$$

In the following we shall be using this relation, not for yield and effort, but for quantities proportional to yield, effort and CPUE, so that the final results is a measure proportional to effort. Therefore we call it "relative effort".

In the example given in Table 9.8.1 we consider the effort of four gears measured in number of units per year. As can be seen, the original effort units are not compatible. The yield corresponding to the effort is also available from a sampling scheme. The four sampled gears given in the table are assumed to constitute only a part of the total catch of the stock in question. To make the different gear types (effort units) compatible each unit must be converted into CPUE, which then in turn is converted into relative CPUE as shown in the table. The relative catch per unit of effort of year y is defined:

$$\text{Relative CPUE for year } y = \frac{\text{CPUE}(y)}{\overline{\text{CPUE}}(y1,y2)}$$

$$\text{where } \overline{\text{CPUE}}(y1,y2) = \frac{1}{y2-y1} \sum_{j=y1}^{y2} \text{CPUE}(j)$$

when a time period over the years y1,y1+1,...,y2 is considered. (In Table 9.8.1. y1 = 1970 and y2 = 1975).

As the relative CPUE of a gear has no dimension you may say that we have obtained compatible units of CPUE by the relative CPUE. The relative CPUEs can be added. If all CPUE observations were proportional to the population size (cf. Section 4.3) the relative CPUEs would become identical for all gears as in that case:

$$Yi(y)/fi(y) = qi * \bar{B}(y) \quad i = 1,2,3,4$$

where qi is the catchability coefficient specific for gear no. i. The relative CPUE would then become independent of qi, as (compare Table 9.8.1):

$$\text{Relative CPUE} = \frac{Yi(y)/fi(y)}{0.25 * \sum_j Yi(j)/fi(j)} = \frac{\bar{B}(y)}{0.25 * \sum_j B(j)}$$

In that case we would from:

$$\frac{\text{Total yield}}{\text{Relative CPUE}} \text{ is proportional to total effort} \qquad (9.8.1)$$

get a measure proportional to the total effort, as the relative CPUE is proportional to CPUE. We usually shall have to add the relative CPUEs and use their mean value as the denominator in Eq. 9.8.1.

Some gears are less important than others. The purse seine in Table 9.8.1 is the dominating one as far as development in yield is concerned, whereas the pole and line fishery is rather unimportant in terms of yield. This is accounted for by calculating the sum of relative CPUEs weighted by the corresponding yields. For example for year 1970 (see Table 9.8.1):

$$\frac{100 * 1.089 + 300 * 1.047 + 40 * 1.11 + 350 * 1.04}{100 + 300 + 40 + 350} = 1.052$$

Dividing the total yield (including the gears not covered by the catch/effort sampling scheme) by the weighted sum of relative CPUEs then gives a figure proportional to the total effort. These figures are shown in the column "YT(y)/R(y)" in Table 9.8.1, whereas the last column contains the normalized relative effort. This concept is introduced in order not to confound relative effort with absolute effort.

The first gear in the table, purse seine, is the important one, in the sense that the trend for this gear is the same as the trend for the combined relative effort. The beach seine shows the opposite trend, but because it is a relatively unimportant gear it has less influence on the combined effort. If the purse seine catches had been only 10% of what they were the trend would be changed as the beach seine then would become a relatively more important gear, as shown in Table 9.8.2.

The method described above has been used by the North Sea Round Fish Working Group of ICES (ICES, 1980). This method does not require a direct comparison of the different boat types. It requires only a type of data which is often available. The method can be questioned, and in fact, the results achieved by the ICES working group were not "overwhelmingly convincing" when the ICES working group correlated the relative effort figures with the fishing mortalities obtained from VPA.

A more direct (and probably more dependable) method to standardize effort is that suggested by Robson (1966) (discussed in Gulland, 1983). It does, however, require additional data. The method works with the concept of "relative fishing power". With the fishing power of vessel B relative to vessel A we mean:

$$PA(B) = \frac{\text{CPUE of vessel B}}{\text{CPUE of vessel A}} \qquad (9.8.2)$$

when the two boats are fishing under the same conditions (at the same time and in the same area). Often vessel A is called the "standard vessel".

Table 9.8.2 **Exploring the concept of relative effort for several gears combined**

year	Table 9.8.1 E(y)	If purse seine catches were only 10% of those given in Table 9.8.1
1971	0.653	1.014
1972	0.734	1.018
1973	0.903	1.035
1974	1.146	0.984
1975	1.564	0.949

Suppose the boats participating in a certain fishery can be divided into a number of homogeneous groups, so that each group consists of boats with the same fishing power. Suppose that the CPUE is in units of catch per time unit (e.g. catch per trawling hour). Suppose that there are 5 groups of boats in the fishery, and that the following data have been collected:

Boat type	A(standard)	B	C	D	E
Fishing power	1.0	PA(B)	PA(C)	PA(D)	PA(E)
Number of boats	NA	NB	NC	ND	NE
Average number of fishing days per boat	dA	dB	dC	dD	dE

The total effort would then be estimated by:

$$\text{total effort} = \qquad (9.8.3)$$

$$NA*dA + PA(B) * NB*dB + PA(C) * NC*dC + PA(D) * ND*dD + PA(E) * NE*dE$$

In certain cases it can be assumed that the fishing power is proportional to some characteristics of the boat or gears which are relatively easy to obtain, such as GRT (tonnage) or HP (horse power) or their product for trawlers and, for example, the number or length of nets for gill netters. As we are usually only interested in the relative effort, the PX's in Eq. 9.8.3 can simply be replaced by the boat/gear characteristics.

9.9 The Deriso/Schnute delay difference model

A family of models which attempts to be a compromise between the surplus production models and the age-structured models has been presented by (for example) Deriso (1980), Ludwig and Walters (1985), Ludwig (1987) and Schnute (1985, 1987).

These models, however, have limited use as they were developed for long lived (slow growing) species, which are not exploited in the first part of their life. They are based on a series of rather strong assumptions which make them inapplicable to certain species. Schnute (1985) states about the models: "Among other things, they reflect the undeniable reality that the population consists of cohorts that get one year older each year". This clearly shows that these models are not intended for stocks where this "reality" can be denied, e.g. shrimps. Actually, the methods give nonsense results for such species (e.g. negative biomass).

The theoretical biological basis for these models does not go beyond the models already introduced (the basic reasoning is similar to that of Pope's cohort analysis), but the mathematical theory applied for the estimation procedure is somewhat more sophisticated than for most of the models presented in this manual. The main difference from the age-structured models given in this manual is that the Ford-Walford equation replaces the von Bertalanffy model for growth in weight (Eq. 3.3.6.1).

The models are highly sophisticated and not quite simple to use. Their description in the language of this manual would require a large chapter indeed. The interested reader is therefore referred to the above mentioned papers by Deriso, Schnute and others.

10 MULTISPECIES/MULTIFLEET PROBLEMS

So far, the models and methods described have dealt mostly with a single stock exploited by one fleet. However, this situation is the exception rather than the rule. In most cases a fleet exploits several stocks and several fleets compete in exploiting the same resources. In this connection we operate with three main types of interaction between components of a multispecies/multifleet system:

1. Biological interaction
2. Economic interaction
3. Technical interaction

Biological interaction means the interaction between fish stocks (and within fish stocks) caused by predation and food competition.

Economic interaction means the competition between fleets (e.g. between "industrial fishery" and "artisanal fishery"). The more one fleet catches of the limited resource the less will be left for its competitors.

Technical interaction means that the fishery on one stock creates fishing mortality on other stocks because the fishery is either a multispecies fishery or because of inevitable bycatches.

This chapter presents a brief discussion of some aspects of the above mentioned three kinds of interaction. It is not intended to explain all the models at a level which enables the reader to apply them in practice. One good reason is that many aspects of the models are still not well investigated and not well understood. For example, it is still the case (in 1989) that all fish stocks in the ICES area (North East Atlantic) are assessed by means of single species models.

Various approaches to models taking interactions into account have been suggested during the last decade. The majority of the models are extensions of the single species/single fleet models presented in the foregoing chapters, so that the theory of the simple systems is a necessary background to multispecies/multifleet theory. Multispecies/multifleet assessment with special reference to tropical fish stocks is reviewed in e.g. FAO (1978), Pope (1979, 1980), Saila and Roedel (1980) and Pauly and Murphy (1982).

10.1 Surplus production models applied to a multispecies/multifleet system

The simplest way to deal with the multispecies/multifleet system is to apply the surplus production models (cf. Chapter 9) to the total catch of all species and the total effort by all fleets. Fig. 10.1.1 shows a plot of the total yield of the Gulf of Thailand trawl fishery against total trawl effort (from Pauly, 1984, after SCSP, 1978). Applying for example the Schaefer model (Eqs. 9.1.2 and 9.1.5) to the yield of all species caught by all fleets would give an estimate of the total MSY for the sea area in question. This approach, however, combines so many complex interactions in such a simple model that its general applicability can be questioned. For example, the surplus production model assumes that the curve is reversible which cannot be true if each stock follows the Schaefer model. This aspect is illustrated in Fig. 10.1.2 which shows a (hypothetical) three species fishery. After effort level F1 (see figure) species A is eradicated and after level F2 species B is gone. Thus, going back in effort from, say, level F3 would according to the Schaefer model, be along the curve for species C. Further, if we assume each stock to conform to the Schaefer model, the total yield of all stocks may not be a parabola. The descending part of the total yield curve will rather be something like

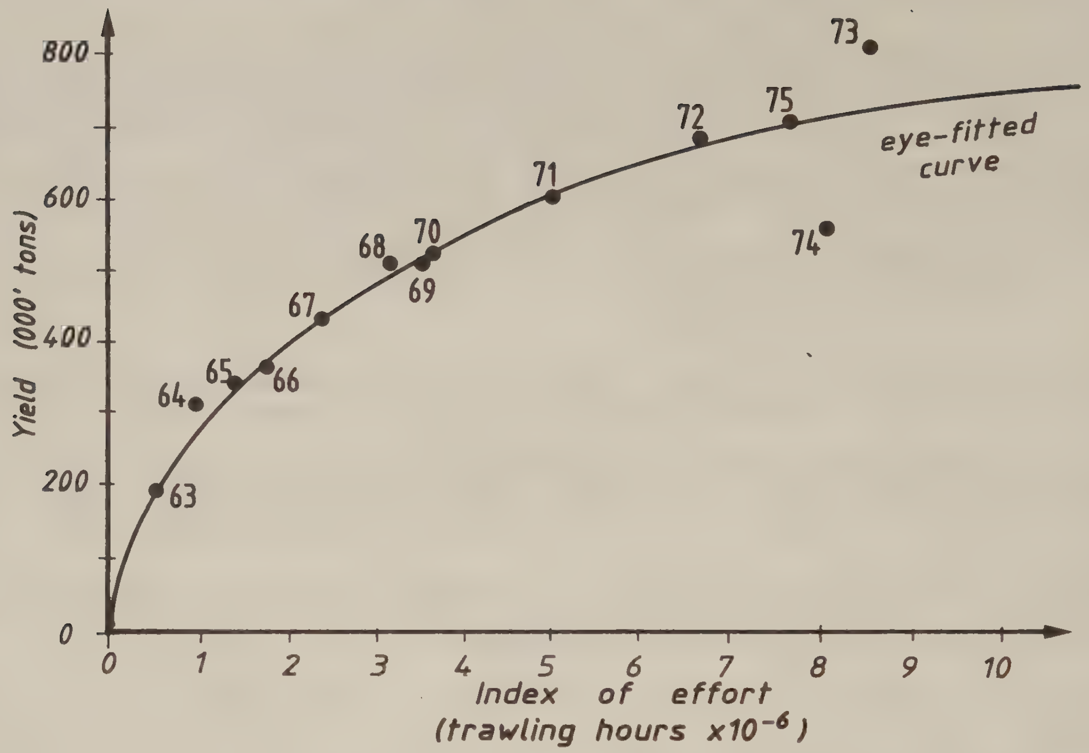


Fig. 10.1.1 Catch and effort from the Gulf of Thailand trawl fishery (from Pauly, 1984. Data derived from SCSP, 1978)

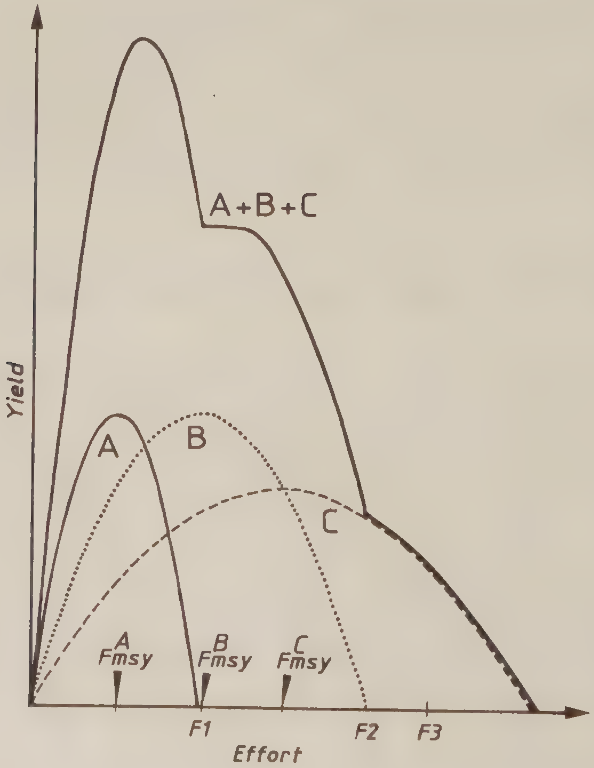


Fig. 10.1.2 Illustration of the changes of yields in a three species system in the simple Schaefer model

an exponential decay curve, as illustrated in Fig. 10.1.2.

In practice, however, a different picture is often observed as in Fig. 10.1.1 in which a steady increase in effort does not produce a drop in total yield. The curve appears to have a continued upwards trend. This probably would come to an end at some high effort level. Yet, it is difficult at the moment to find examples in tropical demersal fisheries of a collapse of the total yield from all species. For a demersal multispecies fishery we expect the following changes together with the effort increase (Pauly, 1984):

1. A decrease (perhaps extinction) of very large fish, which are easy to catch, old and slow-growing

2. A decrease in the average size of fish caught

3. An increase in the relative contribution of low-value small-sized fish. The removal of the large predators on the food fishes results in a reduced natural mortality

4. Increase of previously insignificant components of the system (e.g. squids or jellyfish) which is also explained by removal of predators and food competitors

Thus, the interpretation of results from the simple surplus production model applied to a multispecies fishery is not straightforward. For a further discussion of these aspects, see e.g. Caddy (1980), Marten and Polovina (1982), Ursin (1982) and Pauly (1984). The surplus production models can, however, be extended so that species interaction is explicitly accounted for in the model as will be discussed in the next section.

10.2 Biological interaction

Several authors have more or less successfully extended single species assessment models to cover biological interactions between species. This is done by introducing terms for mutual predation and food competition. The most promising of these models probably is the "Multispecies VPA" (Helgason and Gislason, 1979; Gislason and Helgason, 1985; Pope, 1979a; ICES, 1984, 1986, 1987 and Gislason and Sparre, 1987). Essentially, it consists of a number of parallel VPAs, one for each fish stock in the sea area considered. The second of the two VPA equations (Eqs. 5.1.5 and 5.1.6):

$C = F\,\bar{N}$ (number caught) and $D = M\,\bar{N}$ (number of natural deaths) is extended to:

$$D = (M1 + M2)\,\bar{N}$$

where M2 is the "predation mortality" and M1 the "residual natural mortality" (i.e. natural mortality caused be diseases, starvation, old age, etc.). Predation mortality for prey species no. i, M2(i) is in principle derived from the following expression (Anderson and Ursin, 1977):

$$M2(i) = (1/\bar{N}(i)) \sum_{j(\text{pred.})} \left[\begin{array}{l}\text{predator j's consump-}\\\text{tion of prey no. i}\end{array}\right]$$

where M1 has to be (more or less) guessed, as the total natural mortality had to be in ordinary VPA.

The consumption of prey by predators is estimated from data on stomach contents of predators (Sparre, 1980 and Gislason and Sparre, 1987). In addition to catch data, the multispecies VPA also requires stomach content data and data on food requirements (from feeding experiments) as input. The multispecies VPA was tested in ICES for the North Sea for the first time in 1984 (ICES, 1984; Sparre, 1984; Gislason and Sparre, 1987) but it has not yet been used as the basis for management of the fisheries.

Pope (1980a) developed a version of multispecies VPA based on length data. This method is the extension of Jones' length cohort analysis to multispecies length cohort analysis.

Pope (1979, 1980) also extended the surplus production model (the Schaefer model) to the multispecies case by introducing biological interaction parameters. In the two species case Pope's model reads (compare to Eq. 9.1.2):

$$\text{prey} \quad : \quad Y1/f1 = a1 + b1 * f1 + c1 * f2$$

$$\text{predator:} \quad Y2/f2 = a2 + b2 * f2 - c2 * f1$$

If the interaction parameters, c1 and c2, have zero value we get two independent Schaefer models. When the interaction parameters are positive ($c1 > 0$ and $c2 > 0$) we can interpret the two species model as a predator/ prey system since the interaction term c1 * f2 produces a higher yield of the prey species when effort on the predator is increased, i.e. when the predators are removed. The interaction term for the predator, c2 * f1, has the opposite effect. If effort on the prey is increased the predator gets short of food and the stock is thus less productive, resulting in a reduction in yield from the predator stock.

10.3 Economic interaction

The economic interaction of several fleets was already introduced with the age based Thompson and Bell model (Section 5.3) to which the reader is referred. Economic interaction can also be described by the Beverton and Holt yield per recruit model (Beverton and Holt, 1957), but we shall not go into that here.

10.4 Technical interaction

Most fisheries are mixed fisheries, i.e. the catch consists of a mixture of different species. Pauly (1984, p. 161) gives a table showing a typical trawl catch from the Java Sea. It contains over 55 species distributed over 29 different families. However, the ten most abundant species in that trawl haul constituted 70 percent of the catch. The most abundant single species constituted 32 percent of the total catch in weight. Disregarding the rare species, we often end up with, say, 5 to 15 important species (important from a commercial and/or an ecological point of view).

10.4.1 A yield per recruit model for mixed fisheries

To illustrate the problem we consider a simple system of two species, A and B. Species A is a large slow growing species and B is small fast growing species. Species A has a low natural mortality and species B a high one. The typical shapes of Y/R curves for species with the characteristics of species A and B are shown in Fig. 10.4.1.1. If species A is the target species of the fishery and B is the (inevitable) by-catch then a Y/R-curve for species B has little practical applicability. Management measures have to be directed at the fishery for the target species, A. If there is a simple relationship between the fishing mortality of the two species, for example a linear one:

$$F(A) = k * F(B)$$

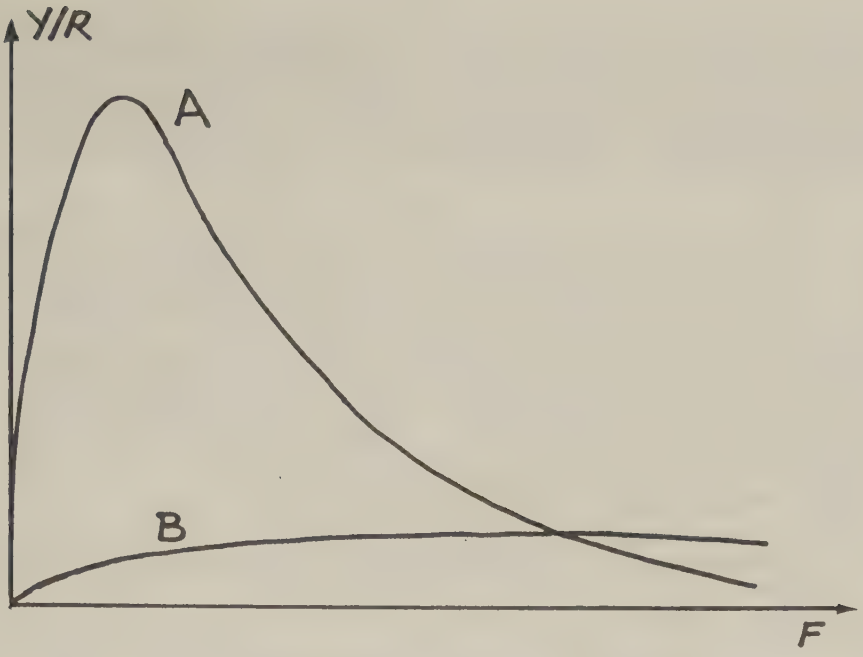


Fig. 10.4.1.1 A: Y/R-curve for a large slow growing species with a low natural mortality

B: Y/R-curve for a small fast growing species with a high natural mortality

where F(A) and F(B) are the fishing mortalities on species A and B respectively and k is a constant, the effects of various fishing strategies on species A can easily be transformed to species B.

Two Y/R-curves cannot be added since they are in different units ("per A recruit" is something different from "per B recruit").

If estimates of relative recruitment are available we can transform the Y/R curves into comparable units (Sparre, 1980; Murawski, 1984). For example, if the number of recruits of species A (on an average) is 0.001 times that of species B we can express the yield for species A in "per recruit" and the yield for species B in "per 1000 recruits", and we can then add the two curves. However, if the prices per kg of the two species are different it may make more sense from a management point of view to multiply by the price before adding. If species A preys on species B the situation becomes more complicated and simple adding of yield curves may lead to erroneous conclusions (Sparre, 1979). To handle this situation a multispecies VPA (Sparre, 1980) might be adequate.

Munro (1983) developed an alternative approach which is less data demanding than the multispecies VPA approach. His model probably represents the simplest way of extending the Beverton and Holt model to take into account species interaction. Murawski (1984) developed an extension of the Beverton and Holt Y/R analysis in which technical interaction in a mixed-species fishery is accounted for.

10.4.2 Assessment of mixed fisheries based on length frequency data

Suppose a certain type of commercial gear catches a mixture of three major species, called A, B and C. One may think of a bottom trawl catching up to 50 different species of which three make up the bulk of the catch. It may well be so that none of these species can be considered the target species.

Thus, the catch consists of a mixture of species which is not determined by the fishing operation but by the availability of the fish. The species in question inhabit the same fishing grounds and are caught together. We assume length frequency data from all three major species to be available. Also, that the kg prices differ between species and between size categories within species.

In this case one cannot treat each species separately and subsequently sum the results in terms of yield. Before a summing makes sense the yield must be converted into units of value. A fishery may for example catch shrimps as the target species and squid as a by-catch. From the point of the fishermen the catch of shrimps is far more important than the catch of squid. Moreover, even if yield is converted into value it is still not possible to sum the results of single species assessments such as the length based Thompson and Bell analyses. It will usually be so that the effort level which for species A gives the maximum sustainable economic yield, MSE, will not be the MSE-level for species B and C.

The approach suggested below combines all species in the estimation of MSE. This assessment of a mixed fishery based on length frequency data works as follows:

Step 1: Do a single species length based cohort analysis on each species separately. This gives estimates of the current fishing pattern for each species, e.g. if species A ranges in length from 5 to 25 cm the F-pattern can be denoted:

FA(5,6),FA(6,7),FA(7,8),.....,FA(25,∞).

Thus, species A is measured in 1 cm classes. If species B ranges from 15 to 100 cm and is measured in 5 cm classes the fishing pattern estimated from cohort analysis may be denoted:

FB(15,20),FB(20,25),FB(25,30),.....,FB(100,∞)

and the results for species C may be

FC(10,12),FC(12,14),FC(14,16),.....,FC(50,∞).

For each species we further get the numbers in the first length group (the recruitment number) from length based cohort analysis.

Step 2: Perform three separate length based yield analyses of the Thompson and Bell type for the three species, as illustrated by Table 5.5.4. Use the same F-factor for each of the three fishing patterns in each prediction. Add the values of the yields of the three species. The results of this exercise is a table similar to Table 5.5.4. This table contains the yield, biomass and value for each species together with the sum of values.

You may, for example, calculate 16 predictions as is done in Table 5.5.4, where each prediction is made as in Table 5.5.3 with the assumptions on fishing patterns listed below:

First prediction:

0,0,...................................0
0,0,...................................0
0,0,...................................0

Second prediction:

0.2 * FA(5,6), 0.2 * FA(6,7),.........,0.2 * FA(25,∞)
0.2 * FB(15,20), 0.2 * FB(20,25),.......,0.2 * FB(100,∞)
0.2 * FC(10,12), 0.2 * FC(12,14),.......,0.2 * FC(50,∞)

Third prediction:

0.4 * FA(5,6), 0.4 * FA(6,7),.........,0.4 * FA(25,∞)
0.4 * FB(15,20), 0.4 * FB(20,25),........,0.4 * FB(100,∞)
0.4 * FC(10,12), 0.4 * FC(12,14),........,0.4 * FC(50,∞)

....
....

Sixteenth prediction:

3.0 * FA(5,6), 3.0 * FA(6,7),.........,3.0 * FA(25,∞)
3.0 * FB(15,20), 3.0 * FB(20,25),........,3.0 * FB(100,∞)
3.0 * FC(10,12), 3.0 * FC(12,14),........,3.0 * FC(50,∞)

Step 3: Use the sum of values to determine the optimum effort level.The assumption behind the method is that when you increase the fishing mortality on species A by, say, 20% the fishing mortality on species B and C will automatically be increased also by 20%.

This approach may be used to assess the combined effect of mesh size changes. If selection ogives are estimated for each species you may use the logistic model (cf. Section 6.1) to determine the fishing pattern. That is, the fishing patterns for the three species are determined by the 50% and 75% retention lengths and the fishing mortality Fm for size classes under full exploitation (cf. Fig. 6.1.3) multiplied by the F-factor, XX. The effect of, say, a 20% increase in mesh size is estimated by calculating three new fishing patterns, using the parameters:

Species A: 1.2 * LA50%, 1.2 * LA75%, XX * FAm
Species B: 1.2 * LB50%, 1.2 * LB75%, XX * FBm
Species C: 1.2 * LC50%, 1.2 * LC75%, XX * FCm

and then do the length based Thompson and Bell analysis for a suitable range of F-factor (XX) values, enabling you to determine the F-level for MSE.

The above suggested computational procedure involves a large number of calculations and it is recommended to use a computer. The LFSA-package (Sparre, 1987) contains the program "MIXFISH" which can perform this assessment of a mixed fishery. The program carries out the single species length based Thompson and Bell analysis for each species as well as the combined assessment and it calculates the MSE in each case, using iteration techniques.

10.4.3 Multifleet mixed fisheries

Section 5.3 discussed the case of two competing fleets (economic interaction) and Section 10.4.2 discussed the case where one type of boat caught several species (technical interaction or mixed fishery). Most fisheries have features of economic interaction and technical interaction as well as biological interaction (Section 10.2). In this section we shall ignore the biological interaction and consider only the combination of economic interaction and technical interaction.

Table 10.4.3.1 shows an example of input data for a multifleet mixed fishery. We are considering seven different pelagic species and four different gears (fleets). Here we assume that only these four gears (fleets) exploit the seven stocks considered. The basic data are numbers caught by size group (or age group):

Table 10.4.3.1 Example of input data for multifleet mixed fisheries for a single time period (e.g. a year). C(y,s,g,i) = number of length class i fish of species s caught by fleet (gear) g during time period y. S = 1,2,...,7. g = a,b,c,d. i = 1,2,...,n(s)

	g \ fleet s \ (gear) species \	a	b	c	d	total catch VPA input
		gill net	long lines	purse seine	pole and line	
1	Yellowfin tuna	C(1,a,i)	C(1,b,i)	C(1,c,i)	C(1,d,i)	C(1,i)
2	Skipjack tuna	C(2,a,i)	C(2,b,i)	C(2,c,i)	C(2,d,i)	C(2,i)
3	Kawakawa	C(3,a,i)	C(3,b,i)	C(3,c,i)	C(3,d,i)	C(3,i)
4	Frigate tuna	C(4,a,i)	C(4,b,i)	C(4,c,i)	C(4,d,i)	C(4,i)
5	Seerfish	C(5,a,i)	C(5,b,i)	C(5,c,i)	C(5,d,i)	C(5,i)
6	Marlin	C(6,a,i)	C(6,b,i)	C(6,c,i)	C(6,d,i)	C(6,i)
7	Sail-fish	C(7,a,i)	C(7,b,i)	C(7,c,i)	C(7,d,i)	C(7,i)

Table 10.4.3.2 Output from the seven individual length based cohort analyses

		population numbers	fishing mortality
1	Yellowfin tuna	$\underline{N(1)}$	$\underline{F(1)}$
2	Skipjack tuna	$\underline{N(2)}$	$\underline{F(2)}$
3	Kawakawa	$\underline{N(3)}$	$\underline{F(3)}$
4	Frigate tuna	$\underline{N(4)}$	$\underline{F(4)}$
5	Seerfish	$\underline{N(5)}$	$\underline{F(5)}$
6	Marlin	$\underline{N(6)}$	$\underline{F(6)}$
7	Sail-fish	$\underline{N(7)}$	$\underline{F(7)}$

$C(y,s,g,i)$ = number of length class (or age group) i fish of species s caught by fleet (gear) g during time period y

$s = 1,2,\ldots,7$ $\qquad$ $g = a,b,c,d$

$i = 1,2,\ldots,n(s)$ $\qquad$ $y = y1,y1+1,\ldots,y2$

Table 10.4.3.1 actually shows only a small fraction of the data. It is just one table in a time series, and each i-index symbolizes a whole length frequency of $i = 1,2,\ldots,n(s)$ observations, where n(s) is the number of length classes in the length range of species s. The rightmost column of the table contains the total number caught:

$$C(y,s,i) = \sum_{g=1}^{4} C(y,s,g,i)$$

= number of length class (or age group) i fish of species s caught by all fleets (gears) during time period y

These numbers by length (or age) group and by time period form the input for s (= number of species) individual cohort analyses or VPAs.

Suppose in the following that we are working with length groups and have performed Jones' length based cohort analysis. We have used as input the average number caught over a time period of several years and have estimated the average number in the population and the overall fishing mortalities for each species. The output of such seven analyses can be summarized as shown in Table 10.4.3.2.

Each entry in Table 10.4.3.2 represents an array (a vector):

$$\underline{N(s)} = (N(s,1),N(s,2),\ldots,N(s,n(s)))$$

$$\underline{F(s)} = (F(s,1),F(s,2),\ldots,F(s,n(s)))$$

where each vector element corresponds to a length group. Each such F of a length group stems from the combined fishing mortality created by all four fleets. The total F for each of the length groups can be re-distributed on the four fleets using:

$$F(s,g,i) = F(s,i) * C(s,g,i)/C(s,i)$$

Table 10.4.3.3 Fishing mortalities partitioned into fleet components estimated from cohort analysis and the numbers caught by each gear

	\ g \ fleet s \ (gear) species \	a	b	c	d	total fishing mort.
		gill net	long lines	purse seine	pole and line	
1	Yellowfin tuna	$\underline{F(1,a)}$	$\underline{F(1,b)}$	$\underline{F(1,c)}$	$\underline{F(1,d)}$	$\underline{F(1)}$
2	Skipjack tuna	$\underline{F(2,a)}$	$\underline{F(2,b)}$	$\underline{F(2,c)}$	$\underline{F(2,d)}$	$\underline{F(2)}$
3	Kawakawa	$\underline{F(3,a)}$	$\underline{F(3,b)}$	$\underline{F(3,c)}$	$\underline{F(3,d)}$	$\underline{F(3)}$
4	Frigate tuna	$\underline{F(4,a)}$	$\underline{F(4,b)}$	$\underline{F(4,c)}$	$\underline{F(4,d)}$	$\underline{F(4)}$
5	Seerfish	$\underline{F(5,a)}$	$\underline{F(5,b)}$	$\underline{F(5,c)}$	$\underline{F(5,d)}$	$\underline{F(5)}$
6	Marlin	$\underline{F(6,a)}$	$\underline{F(6,b)}$	$\underline{F(6,c)}$	$\underline{F(6,d)}$	$\underline{F(6)}$
7	Sail-fish	$\underline{F(7,a)}$	$\underline{F(7,b)}$	$\underline{F(7,c)}$	$\underline{F(7,d)}$	$\underline{F(7)}$

Table 10.4.3.3 illustrates the total fishing mortality partitioned into fishing mortalities by fleet (or by gear). The type of data given in Table 10.4.3.3 together with the estimate of the stock numbers (Table 10.4.3.2) form the input for a length based mixed fisheries Thompson and Bell catch prediction. The first step is to make assumptions on the F-arrays for each fleet (gear), as discussed in Section 10.4.2. This could for example be the assumption of a ten percent increase of the gill net fishery, in which case all arrays in the gill net column should be multiplied by 1.1. Other assumptions could be made for the other gears.

The second step is to sum the fishing mortalities for each species. The third step is to make a length based Thompson and Bell catch prediction on each stock. The fourth step is to distribute the catches by length group between the fleets and convert the catches into values. Finally, the values of catches of different species are summed for each fleet and for the total multifleet fishery.

11 ASSESSMENT OF MIGRATORY STOCKS

Several of the methods presented so far are often insufficient when applied to migratory or schooling fish stocks. They were all based on the assumption that we can take representative random samples of the stock, for example, the whole range of length frequency. When stocks are not vulnerable to fishing, due to horizontal or vertical migrations to areas not covered by the fleet or normal fishing gears during parts of their lifespan, it is usually not possible to sample such stock during these periods. This may simply lead to gaps in the samples for shorter or longer periods, but it is also possible that the samples taken represent different parts of the stock and in such cases it is likely that the data are misinterpreted. In particular, the length-based methods, such as modal progression analysis and catch curve analysis are difficult to apply in the case of migratory fish stocks.

The migrations of fish stocks which have been exposed to fisheries for centuries and to fisheries research for some 100 years (the herring in the North Sea, for example) are sufficiently well known to avoid bias in sampling and misinterpretation of results. The general knowledge of most tropical fish stocks, however, is often very limited and therefore sampling bias, incomplete coverage of the stock and misinterpretation of the data may easily occur.

In this chapter the problems related to migration will be illustrated. Some methods are given which may help in the interpretation of the length-frequency data obtained from migratory fish stocks. These methods should only be considered as first steps to solving complex problems.

11.1 The concept and study of migration

Harden Jones (1968) recognized three types of "migratory movements": drifting with the currents, random locomotory movements and oriented locomotory movements. In a later paper Harden Jones (1984) stated: "I use the word migration in the sense of coming and going with the seasons on a regular basis". Gerking (1953) defined the concept of "homing" as "the return to a place formerly occupied instead of going to other equally probable places". Homing is known to be an important feature of the migration of many fish stocks.

For fish stock assessment purposes, the explanation of why fish migrate is of little importance. The important thing is to know where the fish are at which time of the year.

In this manual we shall look at migration primarily as a source of bias. Migration is here defined as any systematic type of movement of individuals belonging to a stock (cf. the definition of bias given in Section 2.7). Random movements are not considered migrations in the present context because we are only interested in such types of migration that create bias for a size frequency sampling programme. If a fish moves at random the movement will not change the relative probability of being sampled. The distinction between random movements and systematic movements is not obvious. The fish might encounter concentrations of food by moving randomly, and by random movement remain there until the food is nearly exhausted. If the movement of the food items is systematic (e.g., determined by current and other systematic oceanographic features) we would consider the above described feeding movement as a migration, although it was composed of random movements.

Migration is also characterized by being predictable, e.g., for some stocks we are able to predict at which time and where high concentrations can be found.

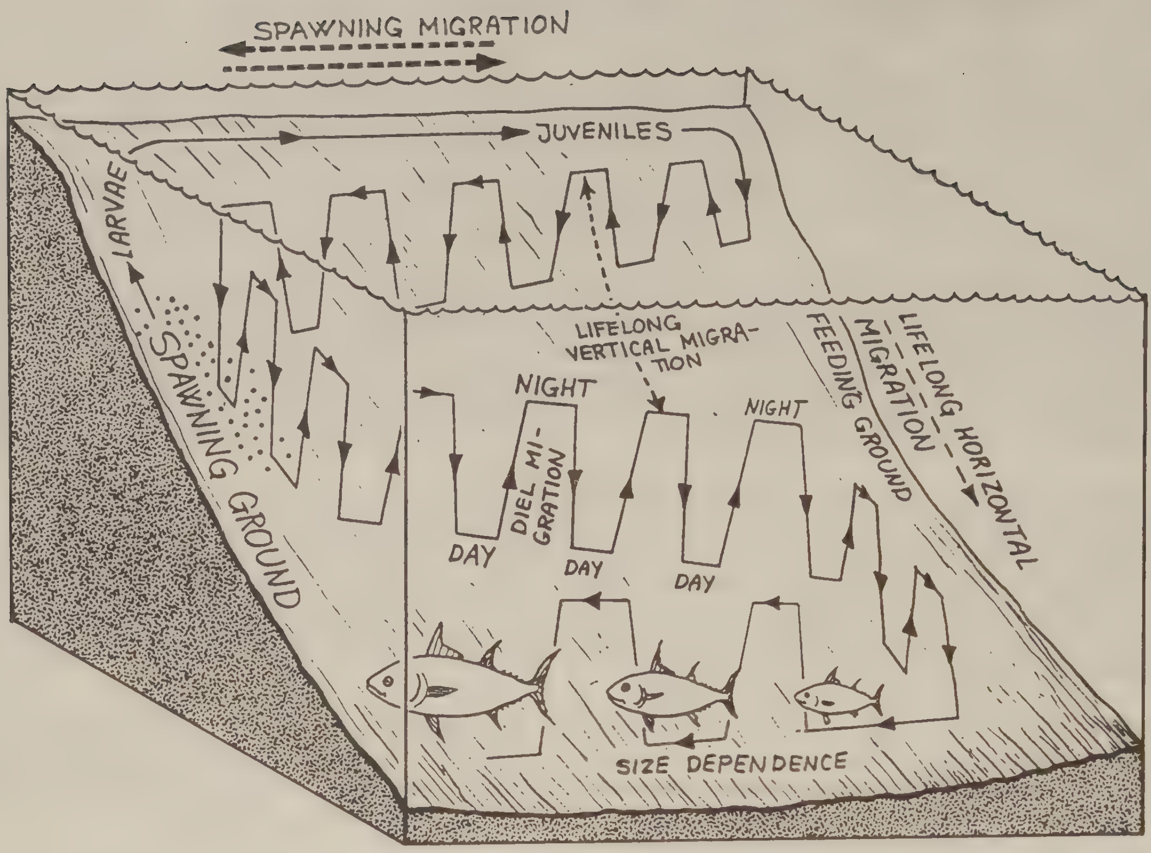

SPAWNING MIGRATION
JUVENILES
LARVAE
SPAWNING GROUND
LIFELONG VERTICAL MIGRA-TION
NIGHT
NIGHT
DIEL MI-GRATION
DAY
DAY
DAY
FEEDING GROUND
LIFELONG HORIZONTAL MIGRATION
SIZE DEPENDENCE

Fig. 11.1.1 **Schematic illustration of five types of potential bias created by migrations**

Most of the migrations causing bias are horizontal movements, along the coast, inshore/offshore or even between rivers/lagoons and marine areas. However, also relatively small vertical migrations and differences in distribution in the water column between different size groups (ages) of the same species may cause bias in our data.

Migrations creating bias may be classified into five main types, of which some are illustrated in Fig. 11.1.1.

1a. Diel vertical migration (e.g., at the bottom during day and in the water column during night).

1b. Diel horizontal migration (e.g., skipjack tuna has been observed to move away and return to a precise location each day (Yuen, 1970).

2. Spawning migration (i.e., annual return to spawning grounds followed by a movement to a feeding ground, or in the case of some cephalopods and penaeid shrimps with only one spawning during their lifespan, migration to the spawning grounds followed by extinction. Homing may be an important feature of this type of migration.

3. Size dependent vertical migration of adults (e.g., skipjack tuna, the younger year classes occur in surface schools, whereas the older specimens move to greater depths).

4. Size dependent horizontal migration (i.e., larger specimens move to deeper waters, while still undertaking the migrations mentioned above).

5. Migration of juveniles. Often juveniles undertake migrations different from those of the adults. The juveniles may remain in the upper layers day and night, whereas the adults are at the bottom during the day. They may also occupy special nursery grounds where the adults are not found. For many species a period in the estuarine or inshore waters is a special phase of life (see for example, Blaber and Blaber, 1980).

An important feature of the migration/bias complex inherent in especially spawning migration is the size dependence. Often, schools with larger average body size arrive earlier to the spawning grounds than the smaller specimens of the same cohort. In addition to the bias problem in fish stock assessment there are other reasons why migration should be studied. One is that of stock sharing between countries (cf. Section 4.1), which is also behind the UNCLOS definition of highly migratory species.

The complex of problems involved in shared stocks contains the problems of stock identification and migration routes. Shared stocks and their management are discussed in, for example, Caddy (1982, 1987) and Caddy and Garcia (1986). Caddy (1982) classifies marine resources based on their movements relative to the EEZs (Exclusive Economic Zones) into five categories:

1. Stocks that lie almost entirely within a single national jurisdiction.

2. Non-migratory resources lying across the boundary between adjacent zones, and which are continuously available in each zone.

3. Migratory species moving across boundary areas only available in each zone on a seasonal basis.

4. High sea stocks that are only occasionally or partially available inside national zones.

5. High sea stocks which occur exclusively outside EEZs.

Especially for categories 3 and 4 the stock identification is a prerequisite for intelligent management of shared stocks. Stock identification and stock assessment has to be based on knowledge of the migration routes of the stock.

There are several ways to acquire knowledge about possible migrations of the species under study. An obvious and cheap way is to tap the knowledge of the fishermen, who surely have noticed seasonal or diurnal variations in the availability of the various target species.

Such fluctuations may also be reflected in the records of landings of various types of fishing boats (statistics) and a general knowledge of the movements of the fleets. Common echosounders are extremely useful for the detection of vertical migrations, while more sophisticated acoustic equipment can be used to map the distribution and estimate the abundance, in particular of small pelagic fish.

The classical way to study movements (and growth) of fish and invertebrates is a tagging programme. Identification tags are attached externally or placed in the body cavity, the fish is measured and released at a known spot, and a reward is given for any tag returned with information on the date and place of capture. Such tagging programmes, if successful, may provide a lot of useful information on net displacements between the point of release and the point of recapture. They do not provide information on what has happened in between those moments and points.

Sophisticated acoustic and radio tags have been developed which allow the continuous observation of the movements of single fish as followed from a research vessel. The latest developments include tags that are released at a preset moment, pop to the surface and transmit to a satellite (applied in tuna research) and tags that record the compass bearing and tilt angle of an individual fish.

In Section 11.5 below some aspects of ordinary tagging programmes are discussed.

11.2 Bias caused by migration

To illustrate the problem, consider a simplified hypothetical fish stock as illustrated in Fig. 11.2.1. This hypothetical fish stock spends one half of the year on the fishing grounds and the other half in an area where it is not exploited and which is also its nursery ground. The stock is composed of two components A and B (you may think of A as the fish recruited before the monsoon and B as the fish recruited after the monsoon). The two components undertake the same migrations but at different times of the year, as shown in Fig. 11.2.1. Suppose the species has a lifespan of only two years and suppose for simplicity that the migrations take place on the 1st of January and the 1st of July. Samples are available only from the fishing grounds.

Assume further that we do not know the migration pattern, but erroneously believe that the entire stock is on the fishing grounds all year round. In this case we are in trouble when sampling data for estimation of growth parameters because the important data for the first half year of life are not available. We may wrongly explain the absence of small fish to be caused by gear selection. We may also observe an apparent negative growth during the period from June to August if we believe that the fish on the fishing grounds in June are the same as those in August. In June component A consists of 1 and 2 year old A-fish, which in July are replaced by 0.5 and 1.5 year old B-fish.

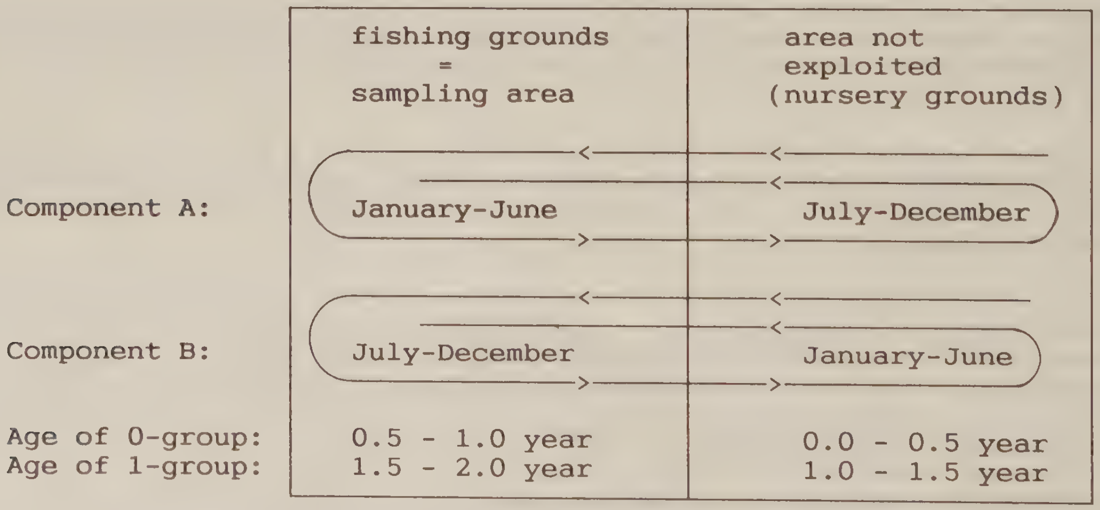


Fig. 11.2.1 Hypothetical simplified system to illustrate the main problem in obtaining random samples from a migratory stock

In this (not very realistic) example one would probably not need to sample the fishery for any long period before one guessed the reason for the missing data for small fish and the apparent negative growth. In reality pelagic fish stocks show a much more complex migratory behaviour and the risk exists that we wrongly interpret phenomena caused by migration as something else. One obvious example is to misinterpret migration as mortality. If, for example, the fish at a certain body size moves to deeper waters where it cannot be caught by the fishing gears it will appear to us (from the samples) as if the fish had died.

The usual situation is that on a given date on given fishing grounds samples can be taken only from a certain fraction of the stock. Also the time of the day and the gear used may restrict the fraction of the stock available to sampling. This section is partly based on the work by Sousa (1988) who discusses sources of bias when assessing stocks of small pelagics, with the scad Decapterus russelli in Mozambique waters as an example. The current knowledge of migration patterns of small pelagics in Mozambique waters is limited, and tagging experiments on small pelagics have not been conducted.

Severe difficulties as for instance apparent negative growth, were encountered when trying modal progression analysis on length frequency samples. Fig. 11.2.2 shows an example of such a problematic time series of length frequencies.

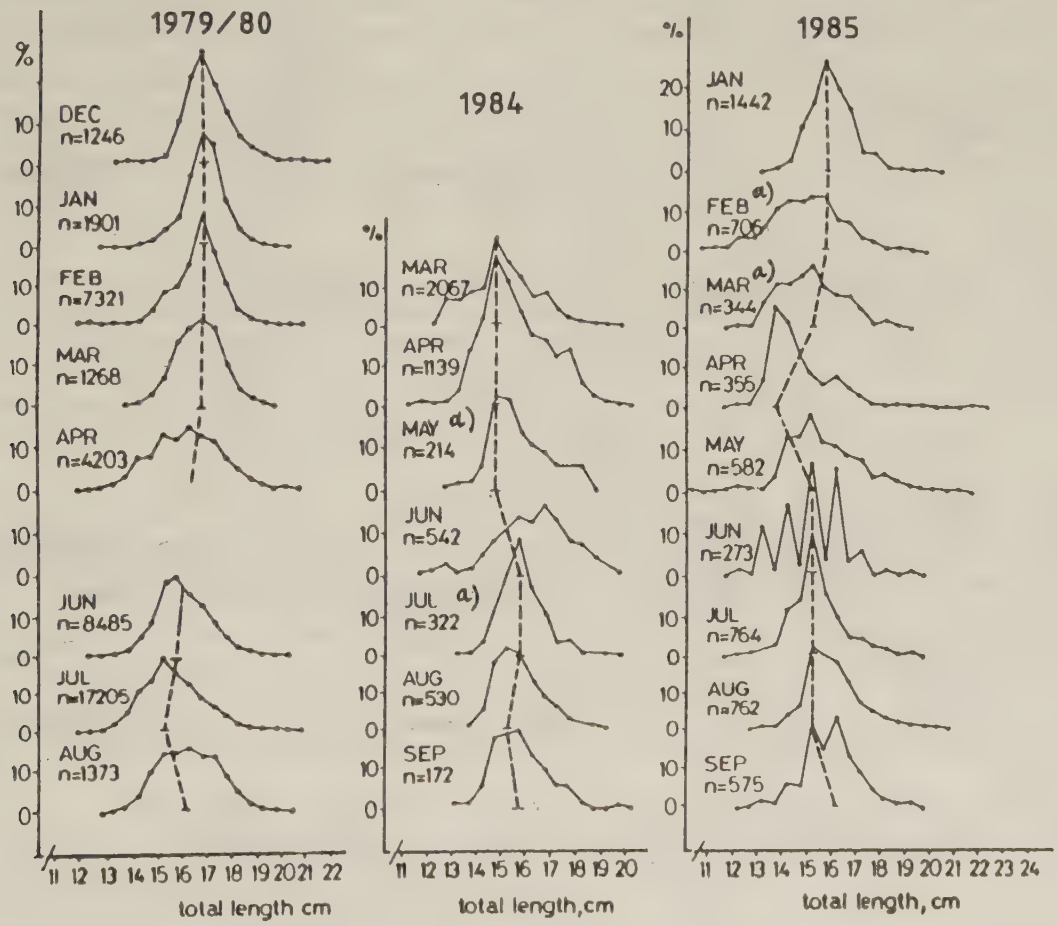


Fig. 11.2.2 Selected series of length frequencies of commercial catches of Decapterus russelli, Sofala Bank, Mozambique, illustrating the problems with the interpretation of modal progression (from Sousa, 1988)

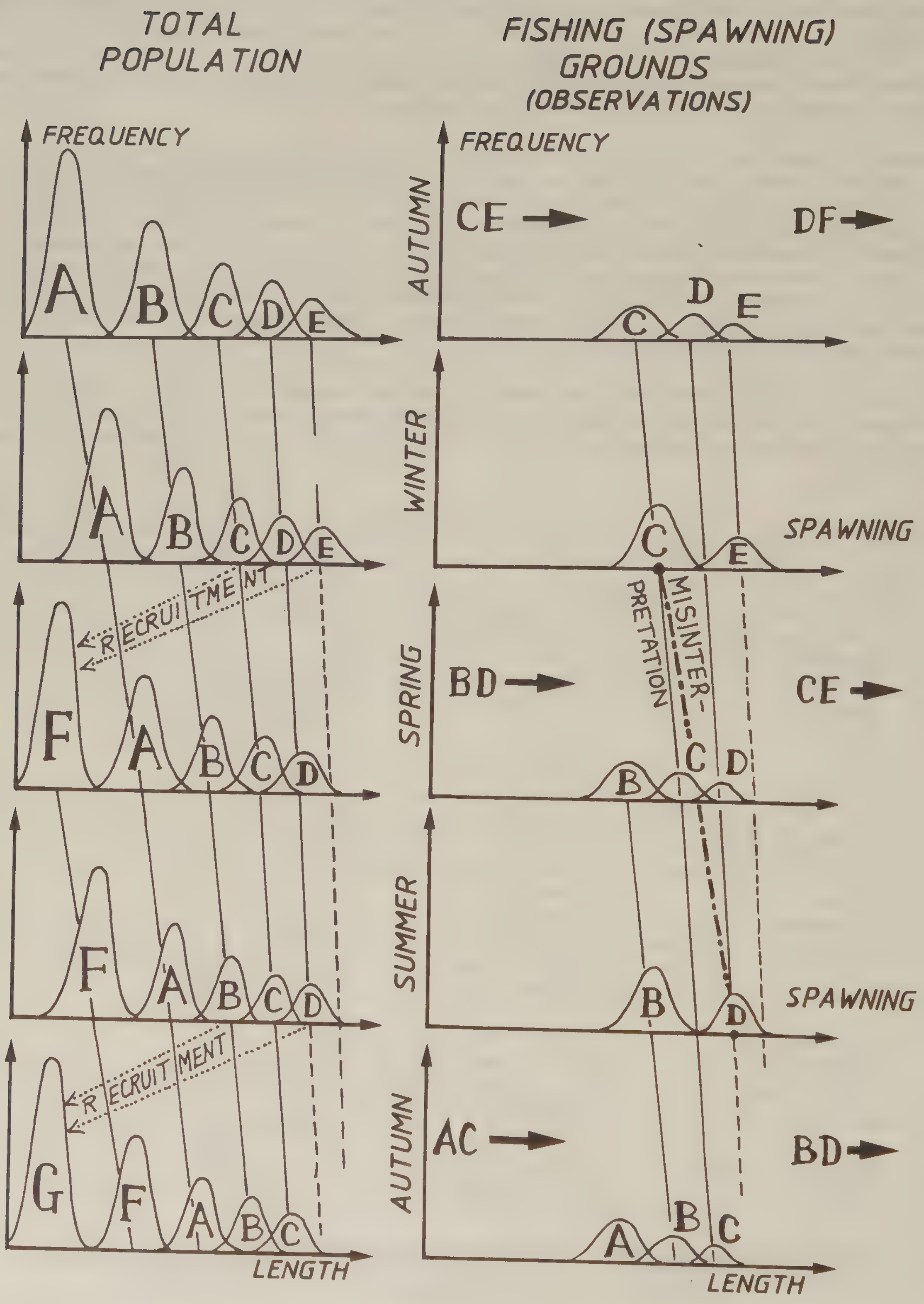

TOTAL POPULATION
FISHING (SPAWNING) GROUNDS (OBSERVATIONS)
FREQUENCY
FREQUENCY
A
B
C
D
E
AUTUMN
CE
DF
C
D
E
A
B
C
D
E
WINTER
C
E
SPAWNING
RECRUITMENT
F
A
B
C
D
SPRING
BD
CE
MISINTERPRETATION
B
C
D
F
A
B
C
D
SUMMER
B
D
SPAWNING
RECRUITMENT
G
F
A
B
C
AUTUMN
AC
BD
A
B
C
LENGTH
LENGTH

Fig. 11.2.3 Schematic illustration of the hypothesis of the migration of small pelagics. One unrealistic simplification is that the cohorts do not overlap. For further discussion see text (from Sousa, 1988)

In the absence of exact data on migration routes, Sousa (1988) searched for plausible explanations of some of the apparent inconsistencies in the time series of length frequencies. The first step was to suggest hypotheses for the migration pattern and the next, to test the hypotheses on the available data.

The reasoning assumes the existence of the migration patterns illustrated in Figs. 11.1.1 and 11.2.3. The hypotheses behind Fig. 11.2.3 is:

> There are two peak recruitments per year. They are denoted: "Autumn cohorts" and "Spring cohorts".
>
> Each year a certain time is spent on the fishing grounds, which are also the spawning grounds. The remaining part of the year the fish are away from the fishing grounds. The spring cohorts are not on the spawning grounds at the same time as the autumn cohorts.
>
> The first half year of life is spent outside the fishing grounds.

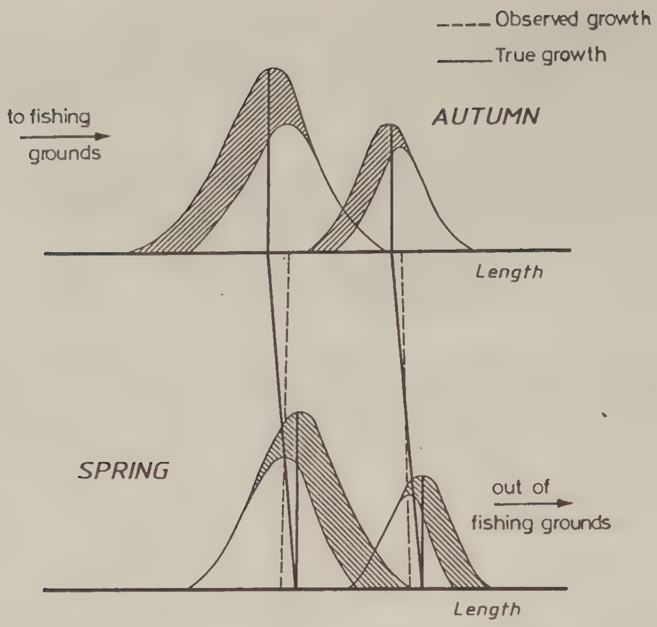
The left hand side of Fig. 11.2.3 illustrates the total population, where A, C and E are the autumn cohorts and B, D and F are the spring cohorts. The right hand side of Fig. 11.2.3 shows that part of the stock which is on the fishing grounds. In the autumn the C and E cohorts are on their way to the fishing grounds and D and F are on their way out. Only a part of C and D has arrived and only a part of E and F has left. In winter C and E only are represented on the fishing grounds but all members of these cohorts are present. The graphs for spring and summer should be explained along similar lines.

This theory could explain why there appears to be so little progression in the length frequency modes. The right hand side further illustrates the risk of connecting peaks from different cohorts. In the simplified Fig. 11.2.3 (with no overlapping between cohorts) it is easy to connect the peaks properly, but with a real data set the overlapping between cohorts would make it much more difficult to do modal progression analysis.

Fig. 11.2.4 illustrates bias created by "size dependent migration" in connection with estimation of growth rates. The example deals with two cohorts. The shaded areas represent the fractions of the two cohorts which are not in the sampling area (fishing grounds). You may say that they are the bias creating part of the stock. The assumption behind Fig. 11.2.4 is that the annual migration (spawning/feeding) is size dependent.

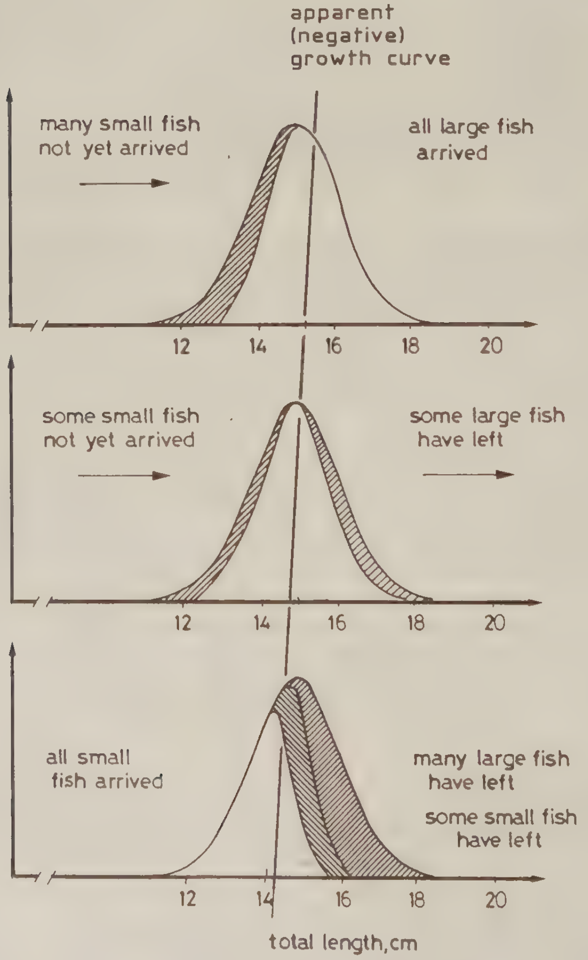
We assume that the cohorts arrive on the fishing grounds during autumn and leave them in the spring. In the autumn sample some of the small specimens (schools with small average size) have not yet arrived on the fishing grounds. The effect will be an overestimate of the average body lengths of the two cohorts. In the spring sample the picture is reversed, all small fish have arrived but some larger fish have left the fishing grounds. Now the effect is an underestimate of the average body lengths. The combined effect of these sources of bias is an underestimation of the growth rate. What we will observe is the growth rate corresponding to the broken line, whereas the full line represents the true average population growth rate.

Fig. 11.2.5 presents a plausible explanation of apparent negative growth along the same lines as Fig. 11.2.4. Thus apparent negative growth may be explained as the combined effect of annual migration and size dependence. Other explanations are possible, however. Apparent negative growth may also be caused by gear selection. If the fishery changes from season to season, e.g. changes geographical ditribution and/or gear the effect may be an underestimation of the growth rate. Selection may in some cases have the opposite effect, i.e. an overestimation of the growth rate.



----- Observed growth
_____ True growth
to fishing grounds
AUTUMN
Length
SPRING
out of fishing grounds
Length

Fig. 11.2.4 Schematic representation of the bias caused by size dependent migration. Dashed zones indicate that part of the population that is not sampled (from Sousa, 1988)



apparent (negative) growth curve
many small fish not yet arrived
all large fish arrived
12
14
16
18
20
some small fish not yet arrived
some large fish have left
12
14
16
18
20
all small fish arrived
many large fish have left
some small fish have left
12
14
16
18
20
total length,cm

Fig. 11.2.5 Schematic representation of possible reasons why apparent negative growth can be observed (from Sousa, 1988)

Fig. 11.2.6 illustrates bias caused by size dependent vertical migration and/or gear selection. Fig. 11.2.6A shows age/length data obtained from otolith readings of 118 D. russelli caught by two different gears, bottom trawl and pelagic trawl (from Sousa, 1988). In this case daily rings were observed so that the age of fish is counted in days. Smaller specimens were caught by pelagic trawl and larger specimens by demersal trawl. The two groups appear to follow different growth curves. This can be explained by size dependence of the lifelong vertical migration. The slower the individuals grow the longer they will remain in the upper layers. If growth curves were estimated for each group of fish the result would be two quite different growth curves. The growth curve shown superimposed on Fig. 11.2.6A is based on all observations. The picture becomes clearer in Fig. 11.2.6B where the residuals (deviations from the estimated growth curve in Fig. 11.2.6A) are plotted against age. Alternatively, it is possible that the findings presented in Fig. 11.2.6 may be explained by higher swimming speed of larger specimens which reduces the catchability of the pelagic trawl for these size groups.

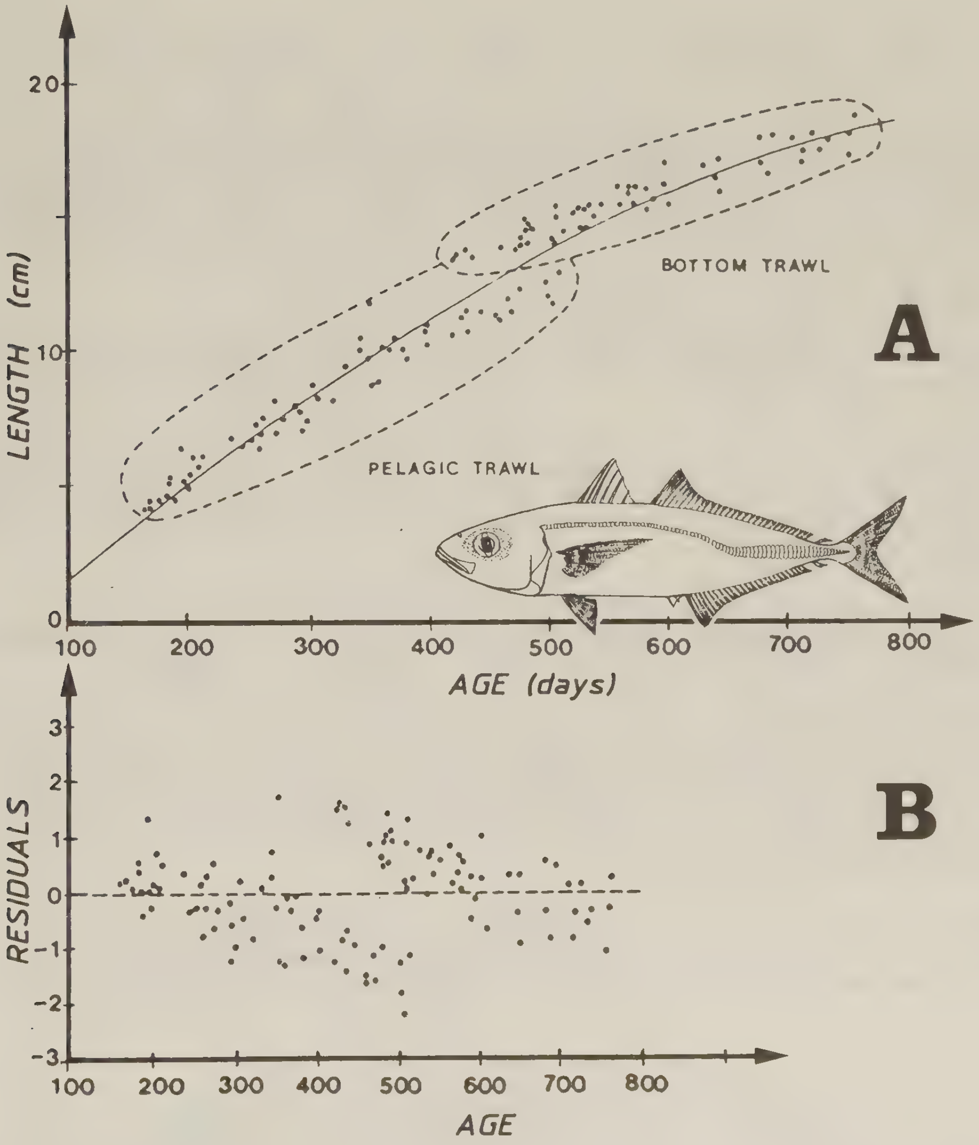


Fig. 11.2.6 Growth curve of Decapterus russelli on Sofala Bank, Mozambique, based on daily age readings of fish sampled by the commercial bottom trawl fishery and by a pelagic trawl survey, both in 1982 (from Sousa, 1988)

11.3 The annual-return matched samples method

The method presented in this section is a simple special case of the "general matched samples method", which will be discussed in the next section. The "matched samples method" is a simple method for the estimation of growth parameters and mortality rates of migratory stocks. As other simple methods it is based on rather strong assumptions. The question is whether these assumptions are reasonable approximations to the reality.

The method is based on the assumption that a fish stock follows a predictable migration route. If this migration route is known (e.g. from tagging experiments) in time and space we are in a position to follow the cohorts and to "match" samples so that they originate from the same cohort.

Consider a simple hypothetical model:

```
Area B(summer) ——>————>————>-- Area C(autumn)
      |                             |

      |                             |
Area A(spring) ——<————<————<—— Area D(winter)
```

A, B, C and D symbolise geographical areas. We assume the stock to undertake the same migration each year and the timing for, say, the spring cohort is as indicated above. To "match" samples in the above model for the spring cohort means to perform the analysis based on samples taken:

In spring in Area A,   in summer in area B,
in autumn in area C and in winter in area D

The samples may originate from one area only, say, A, and in that case only samples collected in the spring (in different years) should be matched.

This model may fit to small pelagics like scads and Indian mackerels and it is known to fit the migration pattern of small pelagics in temperate waters (cf. Section 11.6). It can be used to estimate growth parameters, as described below.

11.3.1 Growth parameters estimated by the annual-return matched samples method

Fig. 11.3.1.1 depicts a series of length frequency samples of a species taken every four months over a 2 years period. Note that the distributions in subsequent periods often indicate a negative growth. Following the model we can connect the modes in the distribution for the corresponding months in subsequent years (e.g. Jan. 1982, Jan. 1983 and Jan. 1984), as in a normal modal progression.

The next step could then be to use the Gulland and Holt plot (cf. Section 3.3.1). We use the "Exact Gulland and Holt equation", (Eq. 3.3.1.2), as the time interval Δt is relatively long, assumed to be one year throughout:

$$L(t+\Delta t) - L(t) = b * L_{\infty} - b * L(t)$$

There are four estimates of $\Delta L/\Delta t$ with Δt = 1 year, namely: (L5-L1), (L6-L2), (L7-L3) and (L8-L4). The mean value of these four observations is 8 cm while the mean of the mean lengths of the first modes (L1+L2+L3+L4)/4 is 4.25 cm. Then we have:

$$L(t+1) - L(t) = 8 \text{ cm}$$

$$b * L_\infty - b * L(t) = b * L_\infty - b * 4.25 \quad \text{or}$$

$$b = 8/(L_\infty - 4.25) \quad \text{or as}$$

$$K = - \ln(1-b) \text{ (cf. Eq. 3.3.1.3) we get}$$

$$K = - \ln(1 - 8/(L_\infty - 4.25))$$

If we use 18 cm as a first rough estimate of L_∞ we get:

$$K = - \ln(1 - 8/(18 - 4.25)) = 0.87$$

Using L_∞ = 18 cm and K = 0.87 per year we obtain the modal progression of Fig. 11.3.1.2 by the ages 0.31 and 1.31 years:

$$L(0.31) = 18 * (1-\exp(-0.87 * 0.31)) = 4.3 \text{ cm}$$
$$L(1.31) = 18 * (1-\exp(-0.87 * 1.31)) = 12.2 \text{ cm}$$

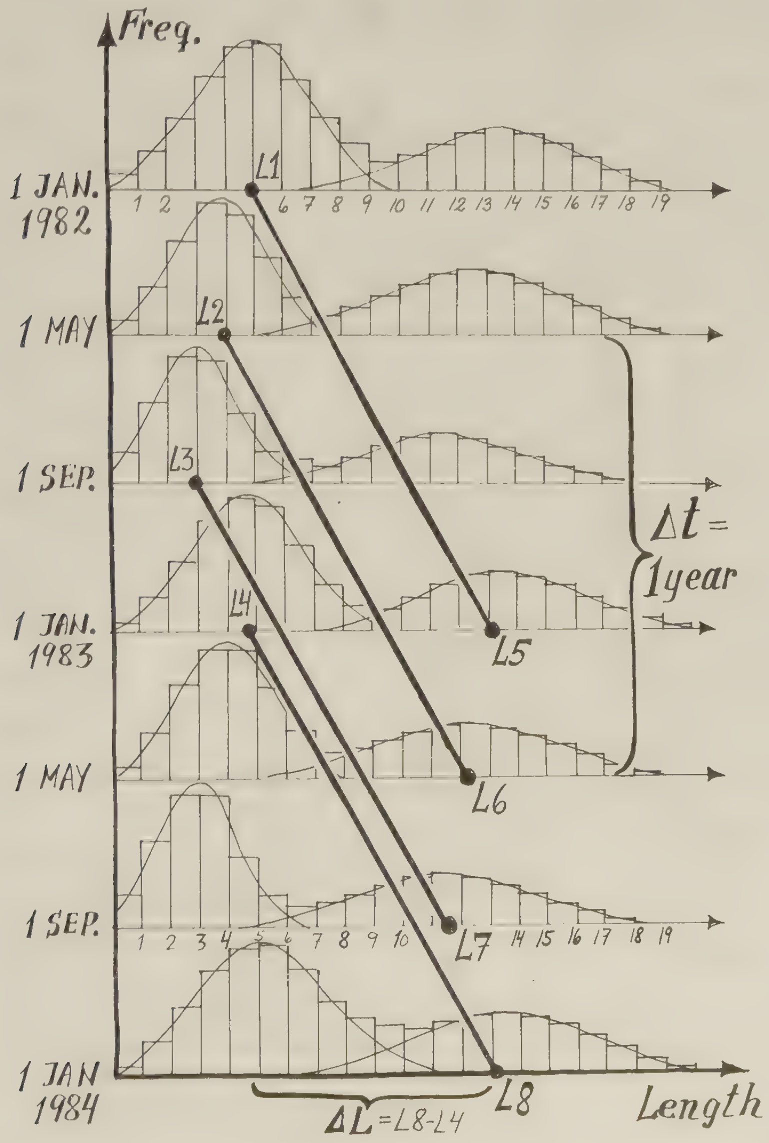

Freq.
1 JAN. 1982
1 MAY
1 SEP.
1 JAN. 1983
1 MAY
1 SEP.
1 JAN 1984
L1
L2
L3
L4
L5
L6
L7
L8
1 2 6 7 8 9 10 11 12 13 14 15 16 17 18 19
1 2 3 4 5 6 7 8 9 10 14 15 16 17 18 19
Δt = 1year
ΔL=L8-L4
Length

Fig. 11.3.1.1 Illustration of the estimate of growth parameters by the annual-return matched samples method. For further details, see text

Table 11.4.1 Hypothetical example to illustrate the "matched samples method"

Area A (number caught per unit of effort)

length	1	2	3	4	5	6	7	8	9	10	total
JAN	0	8465	0	0	0	0	0	0	0	0	8465
FEB	0	0	5732	0	0	0	0	0	0	0	5732
MAR	0	0	0	1758	1153	0	0	0	0	0	2911
APR	0	0	0	0	629	357	0	0	0	0	986
MAY	0	0	0	0	0	125	42	0	0	0	167
JUN	0	0	0	0	0	0	0	0	0	0	0
JUL	0	0	0	0	0	0	0	0	0	0	0
AUG	0	0	0	0	0	0	0	0	0	0	0
SEP	0	0	0	0	0	0	4	15	11	0	30
OCT	0	0	0	0	0	0	0	104	123	32	259
NOV	0	0	0	0	0	0	0	201	355	152	709
DEC	0	0	0	0	0	0	0	178	424	264	867
Total	0	8465	5732	1758	1782	481	47	499	913	448	20125

Area B

length	1	2	3	4	5	6	7	8	9	10	total
JAN	0	0	0	0	0	0	0	0	0	0	0
FEB	0	0	1433	0	0	0	0	0	0	0	1433
MAR	0	0	0	1905	1249	0	0	0	0	0	3154
APR	0	0	0	0	2306	1308	0	0	0	0	3614
MAY	0	0	0	0	0	1870	634	0	0	0	2503
JUN	0	0	0	0	0	271	606	111	0	0	989
JUL	0	0	0	0	0	0	94	74	0	0	167
AUG	0	0	0	0	0	0	50	98	30	0	179
SEP	0	0	0	0	0	0	98	351	239	0	688
OCT	0	0	0	0	0	0	0	422	499	129	1051
NOV	0	0	0	0	0	0	0	212	375	161	748
DEC	0	0	0	0	0	0	0	43	103	64	210
Total	0	0	1433	1905	3555	3450	1482	1311	1247	354	14736

Area C

length	1	2	3	4	5	6	7	8	9	10	total
JAN	0	0	0	0	0	0	0	0	0	0	0
FEB	0	0	0	0	0	0	0	0	0	0	0
MAR	0	0	0	0	0	0	0	0	0	0	0
APR	0	0	0	0	341	193	0	0	0	0	534
MAY	0	0	0	0	0	1252	424	0	0	0	1676
JUN	0	0	0	0	0	738	1650	302	0	0	2690
JUL	0	0	0	0	0	0	1650	1297	0	0	2947
AUG	0	0	0	0	0	0	690	1350	418	0	2457
SEP	0	0	0	0	0	0	216	771	526	0	1513
OCT	0	0	0	0	0	0	0	233	276	71	580
NOV	0	0	0	0	0	0	0	28	49	21	98
DEC	0	0	0	0	0	0	0	0	0	0	0
Total	0	0	0	0	341	2183	4629	3980	1268	92	12494

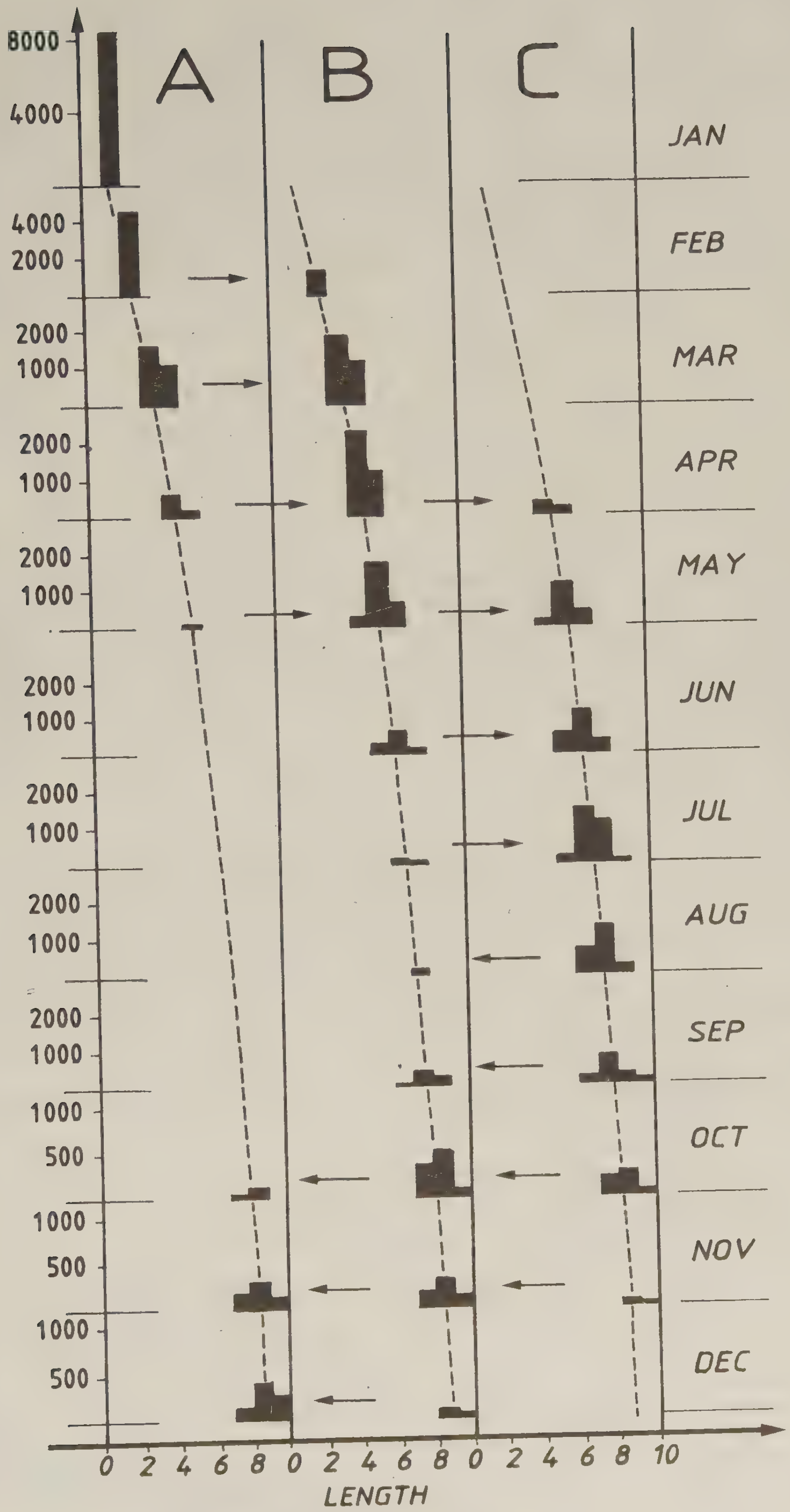


Fig. 11.4.1 Hypothetical example to illustrate the estimation of mortality rates by the "matched samples method". Note that the scales differ from one month to another

11.4 The general matched samples method

This approach assumes that we have knowledge or a hypothesis of the migration route in time and space, and therefore are able to "match" samples so that they originate (or can be hypothesized to originate) from the same cohorts (cf. the beginning of the previous section).

To illustrate the features of the general matched samples method for estimation of total mortalities a simple hypothetical example was constructed. This simplified example deals with one cohort migrating through three areas, A, B and C. From A the cohort moves to B and then to C where it stays for a while; it moves back to B and ends where it started in A.

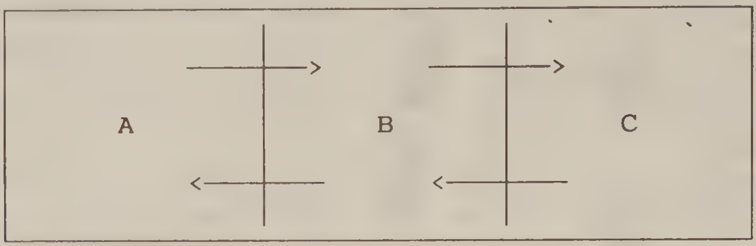


Table 11.4.1 and Fig. 11.4.1 show the example in numbers caught per unit of effort. The curves superimposed on the figure correspond to the growth parameters: L_∞ = 10 cm, K = 2.0 per year and t_0 = 0. The arrows in Fig. 11.4.1 indicate the migration between areas. Ignoring migration we would interpret the first part of the frequencies in area A as a cohort becoming extinct in May and we would estimate the mortality rates as follows (compare Table 11.4.1):

$$Z(JAN) = \ln(8465/5732) = 0.39 \text{ per month}$$

$$Z(FEB) = \ln(5732/2911) = 0.68 \text{ per month}$$

$$Z(MAR) = \ln(2911/986) = 1.08 \text{ per month}$$

$$Z(FEB) = \ln(986/167) = 1.77 \text{ per month}$$

The real mortalities (including the entire cohort in all three areas) are (compare Table 11.4.1):

$$\begin{aligned}
Z(JAN) &= \ln(8465/(5732+1433)) \\
&= \ln(8465/7166) && = 0.17 \text{ per month} \\
Z(FEB) &= \ln((5732+1433)/(2911+3154)) \\
&= \ln(7166/6065) && = 0.17 \text{ per month} \\
Z(MAR) &= \ln((2911+3154)/(986+3614+534)) \\
&= \ln(6065/5134) && = 0.17 \text{ per month} \\
Z(FEB) &= \ln((986+3614+534)/(167+2504+1670)) \\
&= \ln(5134/4346) && = 0.17 \text{ per month}
\end{aligned}$$

The differences between the two sets of Z-values are the "migration coefficients". Fig. 11.4.2 illustrates the difficulties often encountered when trying to estimate growth curves for migratory species. Fig. 11.4.2C shows the migration route for the stock considered. In this hypothetical example we assume that the samples are obtained from the fishing grounds only. The fish are assumed to migrate to and from the fishing grounds depending on their size. During the first year of life they will be available on the fishing grounds when they are at lengths between L1 and L2 and for the second year between lengths L3 and L4.

Figure 11.4.2A shows a time series of unbiased length frequency samples and the growth curves we would estimate from these samples. These samples could be obtained from a sampling scheme which covers the entire distribu-

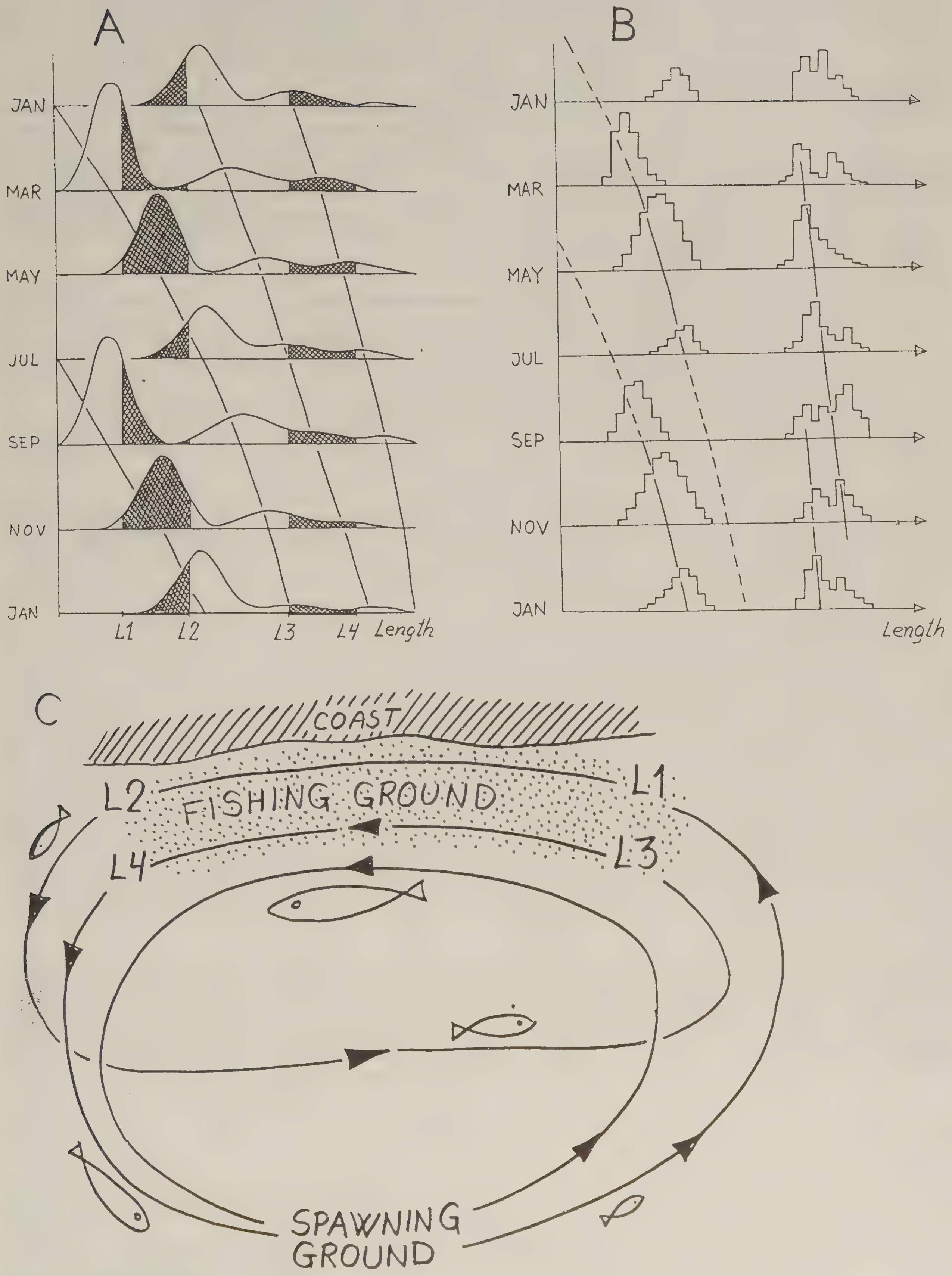

A
JAN
MAR
MAY
JUL
SEP
NOV
JAN
L1
L2
L3
L4
Length
B
JAN
MAR
MAY
JUL
SEP
NOV
JAN
Length
C
COAST
FISHING GROUND
L1
L2
L3
L4
SPAWNING GROUND

Fig. 11.4.2 Illustration of bias problems when estimating the growth curve of a migratory species from "unmatched" samples

tional range in space and time. However, we would not be in a position to obtain such unbiased samples if all samples were taken on the fishing grounds only. If the fish arrived on the fishing grounds exactly at the lengths L1 and L3 and left exactly at the lengths L2 and L4 we would obtain the samples represented by the shaded parts of the frequency distribution on Figure A. In practice we would, of course, not observe such a "knife-edge" recruitment to the fishing grounds. In practice we would observe something like the frequency samples shown in Figure 11.4.2B which show two distinct groups of fish with no overlapping in length frequencies. The group of small fish shows some progression in the modes, whereas the group of larger fish is open for any interpretation. The intermediate length group is absent. This general pattern is often observed when sampling length frequencies from a restricted area, i.e. an area which is only a small part of the distributional range of the stock in question.

The curves superimposed on Figure B are the curves we would obtain from a modal progression. These curves are steeper than those of Figure A (the true growth curves). The modes of the smallest fish of the samples in Figure B show some progression, but less than the progression in the modes of Figure A, i.e. the unbiased samples. The effect of size dependent migration (cf. Section 11.2) is an underestimation of the growth rate. The apparent modes of the large fish in Figure B are difficult to interpret in terms of growth. In practice, these modes may have little to do with growth but rather reflect random noise and/or bias from the sampling procedure. In the case of Figure B the growth curves fitted to the large fish represent an underestimate of the growth rate (compare Figs. A and B), but in practice it might as well produce an overestimate. In general, one should be cautious when interpretting modal progression for length groups approaching the maximum length of the species.

11.5 Assessment based on tagging data

The success of tagging experiments depends on the ability and willingness of the fishermen and others dealing with the catch to report on where and when the marked fish was caught. If the data are used also for estimation of growth parameters the size of the recaptured fish should be reported as well. The fishery must cover a relatively large part of the distribution in space and time of the stock to secure a reasonable number of recaptures for the estimation procedure. It is assumed that the tagged fish constitute a representative sample of the population, and thus have the same basic parameters as the untagged part. Models along these line were suggested by Gulland (1955), Paulik (1963), Seber (1973) and Jones (1977). Seber (1973) presents a comprehensive discussion of the analysis of capture/recapture data.

Kleiber, Argue and Kearney (1983) suggested a model for assessment of Pacific skipjack tuna based on tagging data. This is the traditional catch curve model (Eq. 4.2.7) with modifications to take into account mortality due to tagging, shedding of tags and missing reports on recaptures. The basic equation of the model for estimation of population size, P, and "attrition rate A", (see definition below) reads:

$$r(t) = a * b * N_o * \exp(-t * A) * \frac{C(t)}{P * A} * (\exp(A)-1) \qquad (11.5.1.1)$$

where

1-a = short-term mortality, due to the trauma of being tagged (a = fraction of short time survivors).

A = attrition rate (includes mortalities (F and M); shedding of tags, growth out of vulnerability to the gear and migration out of the area are also considered).

b = fraction of recaptured tags actually returned with usable recapture information.

$C(t)$ = catch in biomass units during time period t.
N_o = number of fish tagged (at time t = 0).
P = standing stock in biomass units (assumed constant in time).
$r(t)$ = number of tag returns during time period t.
t = index of time period.

To see that Eq. 11.5.1.1 is basically the same as the catch equation (Eq. 4.2.7) we introduce:

$N(t) = N_o * \exp(-(t-1) * A)$ = number of marked fish at the beginning of time period t

$F(t) = C(t)/P$, fishing mortality

We assume a = b = 1 and A = Z (no shedding or other type of attrition factors). Eq. 11.5.1.1 then becomes:

$r(t) = N(t) * (F/Z) * \exp(-Z) * (\exp(Z)-1)$ or
$r(t) = N(t) * F(t) * (1-\exp(-Z))/Z$, i.e. the catch equation

The above description is only a short introduction of the basic equation. For a complete description of the model the reader is referred to the original work.

11.6 Estimation of the growth parameters of a migratory stock: the Atlantic mackerel

As mentioned in Section 11.2, seasonally migrating species sometimes migrate earlier in the season the older and bigger the fish are. The problem is analysed in well documented studies of a stock of the North Atlantic mackerel, Scomber scombrus. This stock lives north and west of the United Kingdom and Ireland. The main migration route is shown in Fig. 11.6.1. The stock undertakes an annual migration from the "overwintering area" to the "spawning area", from the spawning area to the "feeding area" and from the feeding area back to the overwintering area. The total distance travelled is in the order of 500-1000 nautical miles per year. Spawning takes place from March to July in the area south of Ireland (see Fig. 11.6.1).

We shall here concentrate on aspects related to size dependent migration as discussed in Sections 11.2 and 11.3. Size dependent migration fo this stock has been demonstrated by, for example, Dawson (1986) and Eltink (1987). The following is based on these two papers, which arrive at similar conclusions from independent sets of observations.

Because it is a species of temperate waters it is relatively easy to read the age of Scomber scombrus from the otoliths. The findings of Dawson (1986) and Eltink (1987) were based on random samples of mackerel caught by commercial fishing vessels as well as research vessels on the spawning grounds during the spawning period. The fish sampled were aged, length was measured, and maturity stages recorded. Samples were taken on a monthly basis. The principal migration routes and the spawning grounds were known beforehand from other investigations (e.g. tagging experiments).

The migration from the overwintering area to the spawning grounds was found to occur in an age-size succession in which the old fish arrive before the young and leave earlier than they do. Even within an age group, the larger specimens arrive before the smaller ones. This is illustrated in Fig. 11.6.2 which shows mean length of fish in maturity stage 6 of random samples collected on the spawning grounds for each month during the spawning season (Eltink, 1987). It shows that the bigger the fish the earlier spawning occurs.

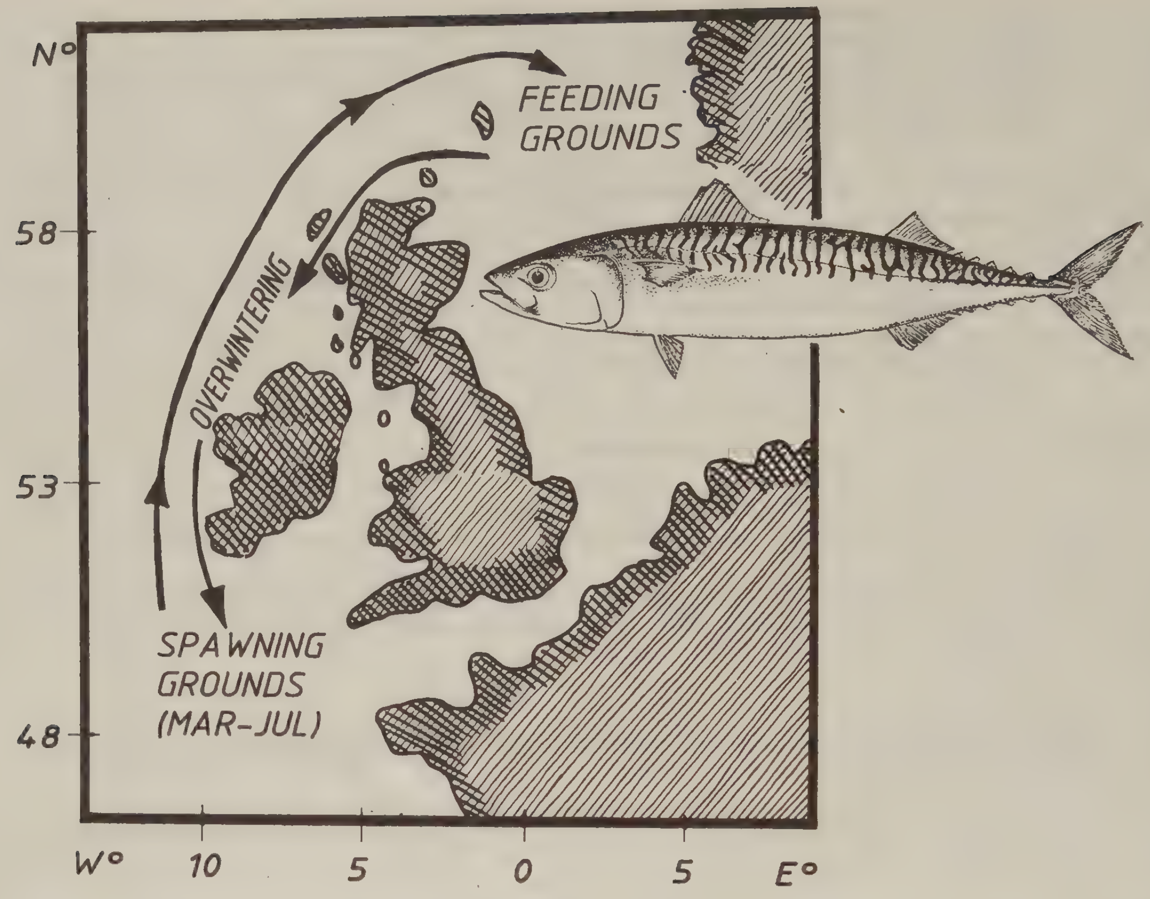

N°
58
53
48
W°
10
5
0
5
E°
FEEDING GROUNDS
OVERWINTERING
SPAWNING GROUNDS (MAR-JUL)

Fig. 11.6.1 The main annual migration route of a stock of Atlantic mackerel (Scomber scombrus) based on various sources including numerous tagging experiments (redrawn from Eltink, 1987)

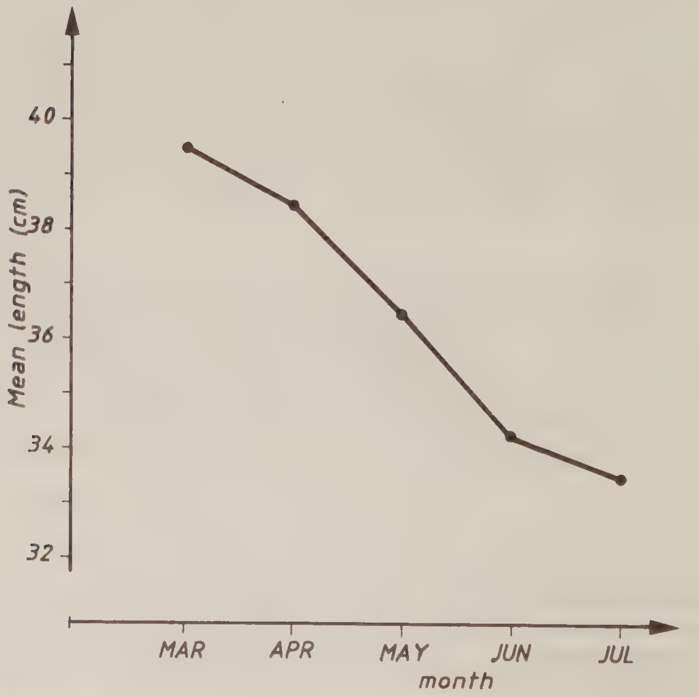


Fig. 11.6.2 Average length of spent (maturity stage 6) Atlantic mackerel (Scomber scombrus) by month. Note that the biggest fish spawn first (redrawn from Eltink, 1987)

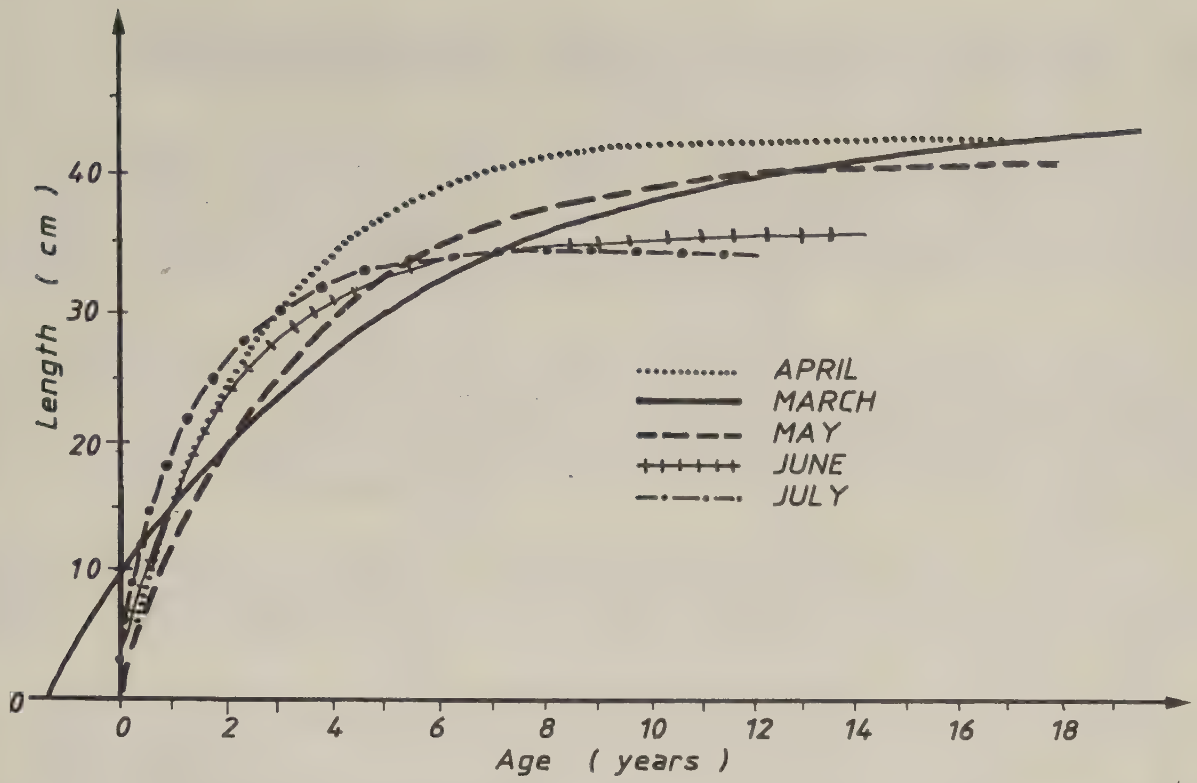


Fig. 11.6.3 Growth curves for Atlantic mackerel (Scomber scombrus) estimated from age/length data of specimens sampled on the spawning grounds from March to July - illustration of a bias problem

Fig. 11.6.3 (from Dawson, 1986) illustrates bias problems when estimating growth parameters for a migratory stock. The five growth curves were each estimated by the matched samples method (cf. Section 11.3). Each curve is estimated from data collected during a particular month on the spawning grounds. As could be expected, the data from the beginning of the spawning period (March) produce a curve for large slow growing fish compared to the data collected at the end of the spawning period, which give a steeper curve with a smaller L_∞.

12 THE STOCK/RECRUITMENT RELATIONSHIP

The stock and recruitment problem may be considered as the search for the relationship between parental stock size and the subsequent recruitment in numbers or the year class strength (cf. Section 8.3). This is a central problem of fish population dynamics, since it represents nature's regulation of population size, whether or not the populations are being exploited.

This chapter presents some considerations on the stock/recruitment relationship problem although it does not present methods to solve actual problems. The reason why the S/R (stock/recruitment) relationship is given this kind of treatment is not that the subject is less important, but rather that really convincing models to handle the problem have not yet been developed.

The following essay on the S/R relationship is based on Beyer and Sparre (1983) and Pauly (1984). Discussions of the S/R relationship with special reference to tropical fish stocks are given in among others, Murphy (1982) and Pauly (1984).

To present the S/R relationship an exceptionally clear example has been chosen (Fig. 12.0.1). It deals with the false trevally (Lactarius lactarius) in the Gulf of Thailand (Pauly, 1980a). In this case there appears to be a well defined relationship between recruitment and spawning stock size. However, as already mentioned in Section 8.3 (Fig. 8.3.2) this is not a typical example of an S/R plot. Fig. 12.0.2 shows the S/R plot for herring (Clupea harengus) in the North Sea for the years 1949-1978. The estimates of stock sizes and recruitment (numbers of one year old fish) are derived from VPA (cf. Section 5.1). This example is more representative for the S/R plots in general.

Obviously, there can be no recruits if no adult fish are left to mature, spawn and produce eggs which hatch and grow to become recruits. The females of most fish species are extremely fecund, producing thousands, even millions of eggs during their adult life. This enormous fecundity has led many fishery biologists to believe that even a very small parental stock should be able to rebuild the stock after each spawning season. It was assumed that features of the abiotic environment (e.g. oceanographic conditions) were the major factors determining how many of the spawned eggs would survive to become recruits. They believe that the spawning stock biomass (biomass of mature fish) was virtually an irrelevant factor for the determination or recruit numbers, except in cases of stock sizes close to zero. This lack of a S/R-relationship was suggested, for example by Beverton and Holt (1957) (see also Beverton, 1963). Later works (e.g. Parrish, 1973 and Saville, 1980) suggest that many fish stocks do display a S/R-relationship. It has been suggested that most of the fish stocks that collapsed in recent decades suffered from recruitment overfishing (cf. Section 8.3) (Murphy, 1966, 1977; Saville, 1980).

However, S/R relationships generally cannot be established directly by plotting the number of recruits (or some index of recruitment) on spawning stock biomass. Rather, it is necessary to simultaneously account for a S/R relationship and the biotic and/or abiotic factors which may affect it. Csirke (1980) demonstrated a strong effect of oceanographic conditions on the recruitment of the Peruvian anchovy.

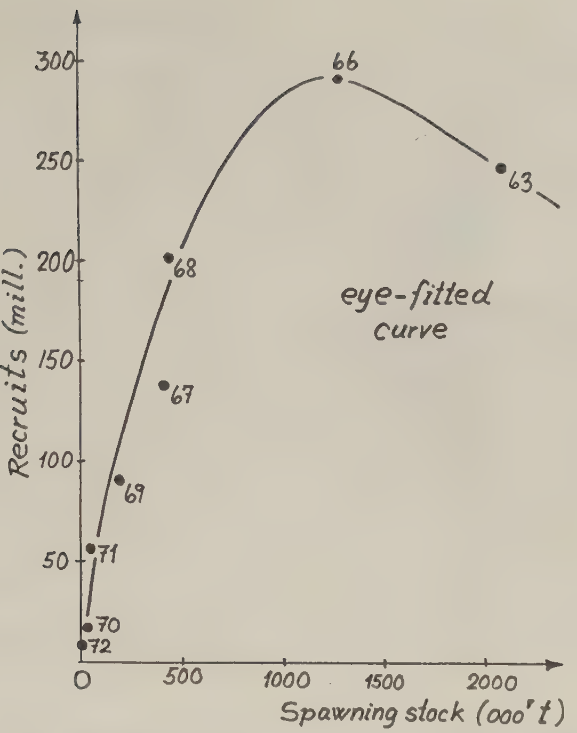


Fig. 12.0.1 Stock and recruitment plot for the false trevally (Lactarius lactarius) in the Gulf of Thailand (from Pauly, 1980a)

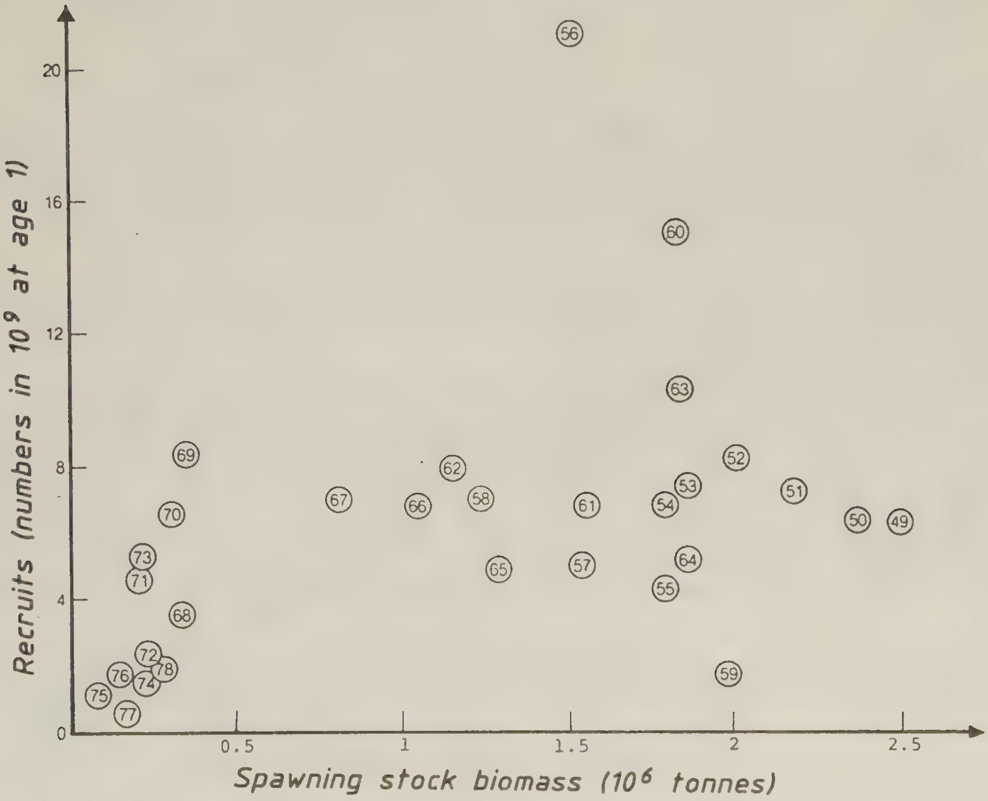


Fig. 12.0.2 Stock and recruitment plot of North Sea herring (Clupea harengus). Number of one year old herring vs spawning stock biomass. Data derived from VPA (reproduced from Beyer and Sparre, 1983)

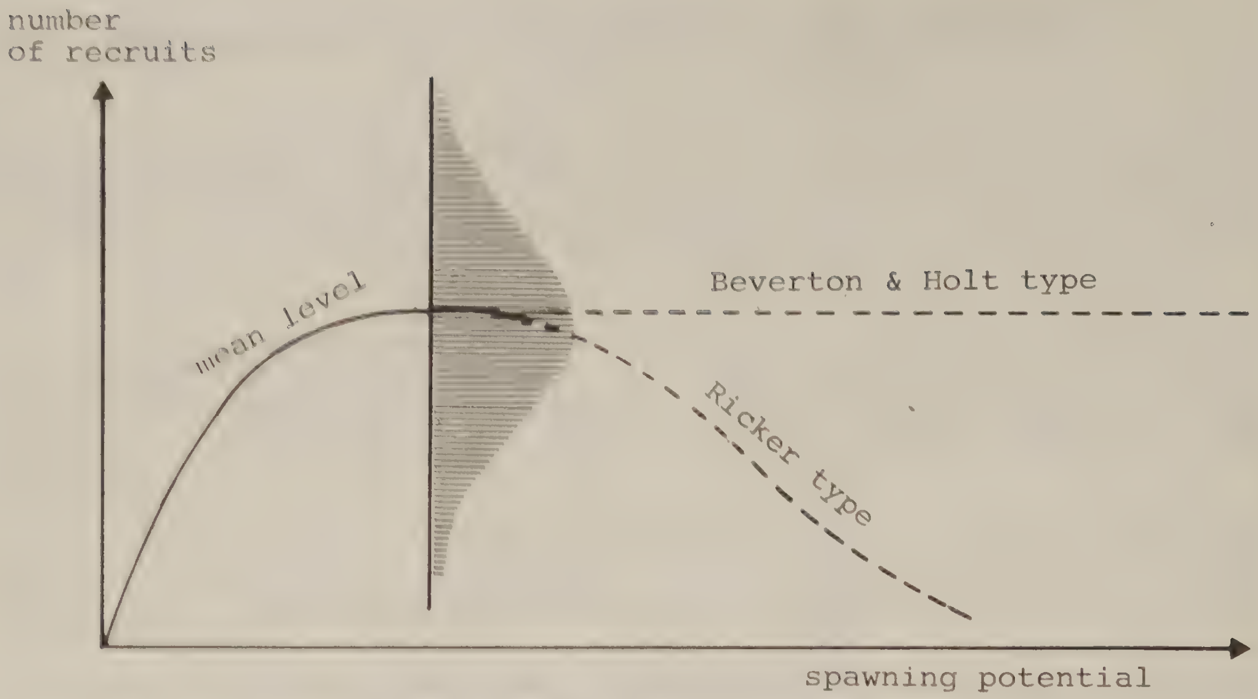

number
of recruits
mean level
Beverton & Holt type
Ricker type
spawning potential

Fig. 12.1.1 The basic form of a recruitment curve. The shaded distribution indicates variation in recruitment about the mean level

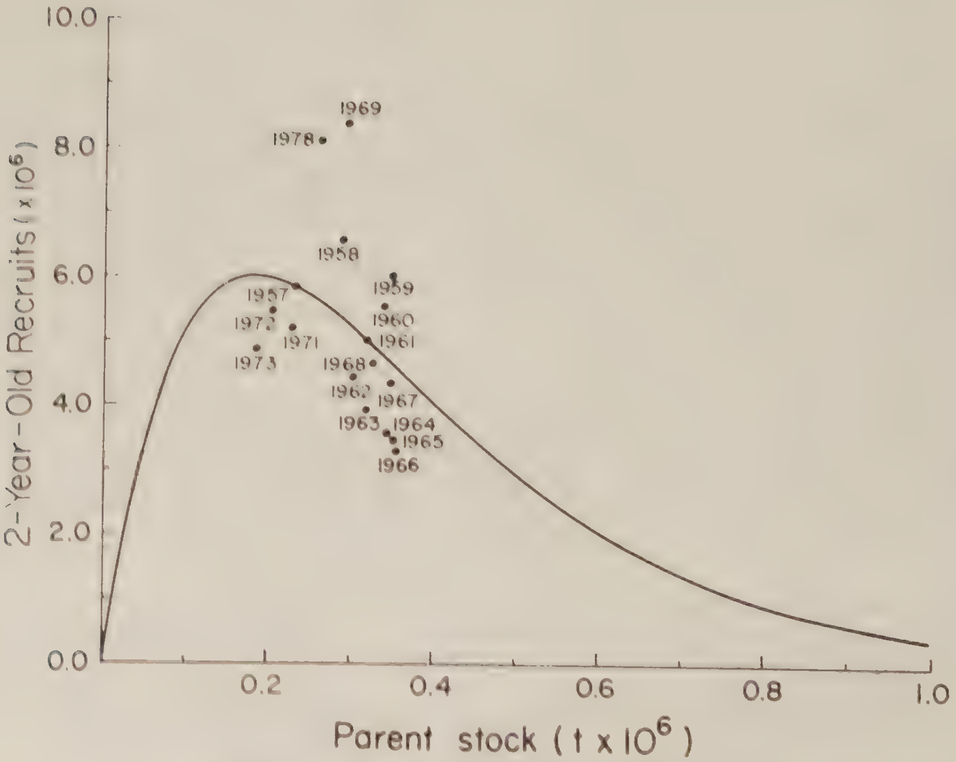


Fig. 12.1.2 Ricker curve fitted to the spawning stock biomass and recruitment of southern bluefin tuna. Data derived from cohort analysis (reproduced from Murphy, 1982, with permission)

12.1 Classical S/R considerations

It has been observed for many fish stocks that the mean recruitment level is almost constant in a large intermediate domain of variation of the parental stock (see Figs. 12.0.2 and 12.1.1). In the case of the North Sea herring (Fig. 12.0.2) it has been estimated that this intermediate range comprises at least an interval corresponding to 10^{12}-10^{13} eggs. The mean recruitment level is about $8 * 10^9$ one year old fish. Thus the probability of an egg developing into a one year old recruit is about 0.0001.

It should be noted that this low rate of survival has nothing to do with fishing, since we are considering mortality in the pre-exploited stage of the life-history (the fish younger than Tr, cf. Section 4.1). The effect of fishing is a reduction of the spawning potential (i.e. growth overfishing, cf. Section 8.2), not of the recruitment. It is only when the rate of fishing is increased beyond a certain level that recruitment is affected (i.e. recruitment overfishing, cf. Section 8.3). In the case of North Sea herring, such a transition from growth overfishing seems to have taken place somewhere in the seventies. But, in principle, we do not know very much about the proper location of this transition zone in the recruitment graph. There is only one thing that is known for certain and that is the point of extinction. We know little about the other extreme of the recruitment curve. Beverton and Holt (1957) suggest the following type of S/R relationship:

$$R = \frac{E}{E + g * Rmax} Rmax \qquad (12.1.1)$$

where g is a parameter and where the number of recruits, R, increases towards an asymptotic level, Rmax, when the egg production, E = (number of females) * (average egg production), increases. This model is derived from a simple density-dependent mortality model:

$$dN(t)/dt = - M(t) N(t) \quad \text{and} \quad M(t) = m1 + m2 * N(t) \qquad (12.1.2)$$

where N(t) is the number of survivors at age t, M(t) the rate of natural mortality of juveniles and m1 and m2 are parameters. Thus, the more survivors there are at age t, the higher their mortality. This mechanism evens out the differences in survival caused by other, biotic or abiotic, factors.

It can be shown (Beverton and Holt, 1957) that Eq. 12.1.1 is the solution to the differential equation 12.1.2.

This is an example of a model describing "density dependent mortality". Density dependent mortality may be explained by food competition. If food is limited the number of starvation deaths becomes a function of the number of specimens which have to share the food.

The characteristic shape of the Beverton and Holt S/R curve is shown in Fig. 12.1.1. The S/R plot of Fig. 12.0.2 is difficult to describe by a curve, but if a curve were fitted it would be of the Beverton and Holt type.

Ricker (1954) suggested a different mathematical model for the S/R-relationship (see Fig. 12.1.1):

$$R = R1 * E * \exp(- R2 * E) \qquad (12.1.3)$$

where R1 and R2 are parameters.

In this model the number of recruits decreases from a maximum level (of R1/(e * R2) when E = 1/R2) towards zero as the production of eggs, E, increases. The recruitment decline is explained by cannibalism of the young by the adults. These matters are described more fully by Ricker (1954, 1975).

The eye-fitted curve in Fig. 12.0.1 is a Ricker type of S/R relationship. However, the descending part of the curve is based on only one point. A few more observations for large spawning stocks may change the picture. Fig. 12.1.2 shows a Ricker curve fitted to data for southern bluefin tuna (Thunnus maccoyii) (Murphy, 1982). This example also illustrates the difficulties often encountered when trying to fit a S/R curve. As Murphy (1982) says about this plot: "there appears to be an underlying density dependent relationship overlaid by considerable environmental variability".

Essentially, the Beverton and Holt model says that above a certain level of spawning stock there is no relationship bwtween parent stock and recruitment, whereas the Ricker model says that this relationship exists for all sizes of the spawning stock, and that there is an optimum spawning stock size (cf. Fig. 12.1.1).

Deriso (1980) and Schnute (1985) suggested a general stock/recruitment model:

$$R = R1 * E\ [1 - R2 * R3 * E]^{1/R3}$$

where R1, R2 and R3 are parameters. For large values of the "shape" parameter R3, the above model reduces to the Ricker model as we then get

$$R = R1 * E \exp(- R2\ E)$$

If R3 = -1 we get the Beverton and Holt model by redefining the parameters to R1 = 1/g and R2 = 1/(g * Rmax) as

$$R = R1 * E/[1 + R2 * E] = (R1/R2)/(E + 1/R2)$$

Most available data on stock and recruitment refer to the medium range of variation in the spawners. It is here that most species appear to maintain a constant mean recruitment level and this constitutes the motivation for the assumption of constant recruitment in classical fish population dynamics. The most remarkable fact, perhaps, is that recruitment shows only relatively small variations about this mean level considering the great reduction in numbers from the egg stage until maturity (Ursin, 1982)

12.2 The stability of recruitment

Table 12.2.1 gives relative recruitment variations for eleven commercially important North Sea species in the period 1963-75 (Ursin, 1982). The table is derived from VPAs from various ICES working group reports. We see for example, that the cod year classes show a factor of six in their variation from the weakest to the strongest.

To consider this variation range in the light of the range of population reduction of cod during the first year of life we may apply the following consideration as an approximation of the situation in the seventies. The mature stock comprises about 200,000 tonnes of cod, half of which are females. Once a year each female spawns 10% of her body weight as eggs, giving a total annual production of 10,000 tonnes of eggs or $2 * 10^{13}$ eggs, since each egg weighs about 0.5 mg wet weight. The mean recruitment level, however, is only $2 * 10^{8}$ one year old fish. Thus on the average only one egg out of 100,000 survives and grows into a one year old cod. The factor of six in recruitment variation implies that the probability of a fish dying (surviving) in the first year of life at most changes from 0.999,997 (.000,003) for a bad year class to 0.999,983 (0.000,017) for a good year class.

Table 12.2.1 **Recruitment variation in North Sea fish stocks 1963-1975. Numbers adjusted to a value of 100 for the most outstanding year class of each species. "Ratio" is the ratio between the numbers in the strongest and the weakest year class (after Ursin, 1982)**

	gadoids					flatfish		clupeids		others	
year class	cod	haddock	whiting	saithe	Norway pout	plaice	sole	herring	sprat	sand eel	mackerel
1963	52	1	14	17	4	100	100	100	-	-	10
1964	49	1	26	23	6	29	21	53	-	-	26
1965	70	2	30	18	0.5	28	11	47	-	-	43
1966	63	12	37	50	7	25	11	66	59	-	62
1967	20	100	100	51	-	21	18	65	62	-	10
1968	19	6	33	55	2	27	9	36	37	-	16
1969	82	2	30	29	-	32	26	78	41	-	100
1970	100	14	33	29	45	25	6	62	20	100	11
1971	18	21	68	30	7	20	14	41	19	21	17
1972	35	4	90	40	16	62	19	19	46	47	4
1973	31	21	63	100	100	40	18	47	91	28	15
1974	51	40	92	27	38	25	7	-	100	86	11
1975	27	9	37	50	18	37	22	-	79	41	4
Ratio	6	100	7	6	200	5	17	5	5	5	25

We could also consider stability in terms of recruitment to the mature stock. During her lifetime a female cod must produce an average of one mature female and one mature male. Only extremely small deviations from this magic number 2 are feasible if the cod stock is to remain at approximately its present level, as it apparently has done for the last 150 years or more. This approximate one-for-one replacement between successive generations holds at least as long as the mature stock is within the medium range of variation as referred to above. Thus, whether we are considering a situation of heavy fishing , where the mature female cod on the average lays about two million eggs in its short life time, or whether we are considering an average production of 40 million eggs per female at a low fishing level, the number 2 still holds.

We do not know what causes this enormous reduction in numbers from the egg stage and leads to the fine adjustment of the number of recruits to the number of mature fish. Note in passing that the North Sea haddock and Norway pout show a remarkable instability in recruitment (see Table 12.2.1). It could be that stabilizing mechanisms do not exist in the North Sea for these two species.

12.3 Towards modeling recruitment

It is difficult to attack the recruitment problem in a sensible way by means of statistical tools only because of the small amount of useful data that in general seems to be available. We only achieve one new data point per species each year. Thus, to explain recruitment variability such as the occurrence of extremely strong year-classes, we may advance a great variety of conflicting hypotheses, all of which lead to different conclusions but none of which can be rejected on the basis of the available data.

The situation is no better with respect to explanations of recruitment stability. It is not possible to discriminate between empirical recruitment models such as those given by Eqs. 12.1.1 and 12.1.3. Neither is it possible beforehand to sort out irrelevant mechanisms of mortality regulation by means of data. Apparently we need to build more biological knowledge on the causes of natural mortality in the first year of life into the recruitment models.

13 DEMERSAL TRAWL SURVEYS

The mean catch (either in weight or in numbers) per unit of effort or per unit of area is an index of the stock abundance (i.e. assumed to be proportional to the abundance). This index may be converted into an absolute measure of biomass using the so-called "swept area method" (Section 13.5).

Bottom trawl surveys are widely used for monitoring demersal stocks when only an index of abundance is required. From unfished stocks (or stocks for which no or few data on the fishery are available) biomass and annual yield estimates may also be derived by bottom trawl surveys. The estimation of total biomass from the catch per unit of effort (or unit area), however, involves several crucial assumptions, leaving such estimates rather imprecise.

Reviews of the theory are given in e.g. Gulland (1975), Saville (1977), Troadec (1980), Doubleday (1980) and Grosslein and Laurec (1982). These reviews also give guidelines for conduct of surveys (planning, design, data collection, data recording, analysis and reporting), some of which are summarized in Sections 13.2-13.4), see also Butler et al. (1986).

This chapter gives first a short description of demersal trawl survey planning and a few guidelines for data recording and for field work (Sections 13.1-13.4). For more detailed descriptions of these subjects the reader is referred to e.g., Alverson and Pereyra (1969, Alverson (1971), Mackett (1973), FAO/UNDP (1975), Gulland (1975), Saville (1977), Flowers (1978), Doubleday (1981), Grosslein and Laurec (1982) and Fogarty (1985). The remainder of the chapter (Sections 13.5-13.8) is a short account of the theory necessary to perform a stock assessment based on trawl survey data.

13.1 The bottom trawl

The bottom trawl (Fig. 13.1.1) is a conical net bag with a wide mouth fitted with weights on the ground-rope and floats on the head-rope (cf. Section 6.3). When the vessel is under way the net is kept open by two otter boards, wooden or iron structures which are towed by the warps attached forward of their center so they tend to diverge. The two otter boards are connected to the net by bridles of rope or wire. These may be up to 200 m long and sweep the sea bed over a wide area. They frighten the fish towards the advancing net and so increase its effectiveness. The shape of net varies depending on the kinds of fish to be caught and on the types of bottom. The ground-rope may be fitted with roller gear (bobbins) so that the trawl can be used on stony bottom without being damaged.

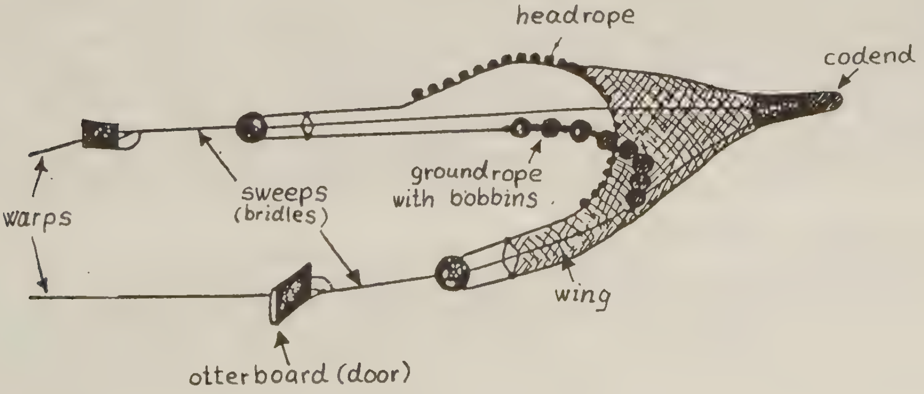


Fig. 13.1.1 Bottom trawl

The tail end of the gear from which the captured fish are removed is called the "codend". This is where most of the size selection takes place. In most cases a relatively small mesh size is required in the codend, in order to obtain a representative sample for the entire size range of the species under investigation (see Section 13.2).

13.2 Planning a demersal trawl survey

Below follows a list of some important items to consider before conducting a survey.

DEFINITION OF OBJECTIVES:

The objective(s) should be specified. Examples of objectives are:

1. Estimation of the total biomass
2. Estimation of the biomass of selected species
3. Collection of biological data for estimation of growth and mortality parameters (e.g. length frequencies)
4. Collection of environmental data

INFORMATION ABOUT THE SURVEY AREA:

Information about depth and bottom conditions to point out trawlable areas and decisions on strata may be obtained from a preliminary survey with echo sounding. Information from local fishermen may be valuable. Information on seasonal winds, currents and migration patterns of fish stocks are important as well.

CHOICE OF GEAR:

The design of the trawl should fit with the expected bottom conditions and the vessel used. If rough bottom is prevalent the gear should be fitted with bobbins in order to avoid damage to the gear. If semi-pelagic species are common a high opening trawl should be used. The mesh size in the codend should be chosen so that the entire size range of the fishable part of the stock is retained by the trawl. Often 10-20 mm stretched mesh is appropriate. The mesh size used for assessment surveys is usually much smaller than the size used by the commercial fisheries.

If different trawls are used parallel or alternate hauls should be carried out in order to estimate correction factors for pooling of data.

SURVEY DESIGN:

A procedure for the selection of stations should be decided on. A fixed grid of stations ensures maximum information on the distribution throughout the area, but not necessarily the most precise estimate of biomass. For estimation of stock sizes a completely randomized design or a stratified random sampling design should be preferred. In most cases a stratified sampling design should be chosen because fish are seldom uniformly distributed and in most cases abundance is related to depth. Stratified sampling often gives a much more precise estimate for the same (or even lower) cost than simple random sampling (cf. Sections 2.8 and 13.7).

ALLOCATION OF HAULS (STRATIFICATION):

Strata are basically constructed in accordance with the density distribution of the fish, so that areas with high/medium/low densities are separated. To do so, some information must be available before the survey. Perhaps the first survey (or the first part of the survey) should have a completely randomized survey design, or the trawl hauls might be evenly spaced. But in the next part of the programme some information on densities and standard deviations of density estimates will be available, and that information can be used for stratification. The standard deviation of catch per unit area (CPUA) is often found to be proportional to the CPUA and will therefore be higher in areas with high abundance.

The optimum allocation of a given number of hauls between strata will be to sample each stratum in proportion to its standard deviation multiplied by the stratum size (cf. Sections 2.8 and 13.7). The distribution of hauls within strata should preferably be at random, but often practical matters dictate the sample design. For example obstacles on the bottom may not allow a proper random distribution of stations.

POSSIBLE NUMBER OF HAULS:

To estimate how many hauls it is possible to make in a given period the following information should be available:

Total number of days available	N
Time spent going to and from the fishing ground (hrs)	t1
Duration of one haul (hrs)	t2
Time used for shooting and hauling the trawl (hrs)	t3
Time to cover distance between stations (average, hrs)	t4
Number of hours available per day (depending on crew, behaviour of investigated species, navigation, etc.	T
Time used for preparations: loading of ice, food, water, repair of gear and equipment (days)	t5

Except for the first and the last day of a cruise, when T should be reduced by t1, the number of hauls per day can be calculated from:

(number of hauls per day) = T/(t2+t3+t4)

(Total number of hauls) = (N-2-t5) * (number of hauls per day)
+ (hauls made first and last day)

It is important to standardize the length of a haul throughout the survey, since the catchability of species and sizes often depends on the duration of the haul. For survey purposes hauls of 0.5 hour or 1 hour are usually the most adequate (see also Section 13.6).

13.3 Data recording

When setting up a plan for a trawl survey a crucial point is to decide on the data items to be recorded, the required precision and how often they should be recorded.

The data items to be recorded are determined by the models by which the objectives of the survey can be achieved, e.g. the swept area method (cf. Section 13.6), length frequency analysis (cf. Chapter 3), mortality estimation (cf. Chapter 4). Such data items would usually include specifications of the gear used, and for each haul the time and position at start and at end of the haul, wire length, wingspread, bottom type and depth. The catch record should include total weight, species composition by weight and length frequencies (for selected species).

The required precision depends on the subsequent use of the data. However, often the precision of a trawl survey is controlled by the number of hauls, which limits our ability to decide on the precision.

It must be known how the data are going to be processed before data recording takes place, particularly in cases where data reduction takes place before the recording stage. Consequently, before you can design appropriate log sheets you must have a rather precise idea of how data will subsequently be processed. When designing a log sheet, practical considerations are important, e.g. where the data are entered, on the bridge, in the laboratory or on the deck.

It is important that the data are well organized to facilitate processing, e.g. by computer. This is especially the case when data from more than one vessel have to be combined. Properly designed recording documents will make the correct recording much easier and more dependable. Several documents are usually required to record all the information collected during a survey. These are in the following categories:

1. A log which summarizes the whole cruise.
2. Details of individual stations (or hauls), the "fishing log", which will generally provide information on the vessel's position, time of start and end of haul, gear rigging etc. Summary information on the catch such as total weight and weight composition by species should also be recorded on the fishing log.
3. Detailed information on the catch. This may be in terms of length, weight, sex, maturity stage etc. for each specimen, or samples of length frequency distributions.

A detailed description of a data processing system for demersal trawl surveys is given in Flowers (1978) who reproduces forms suitable for the work specified as 1-3 above.

13.4 Deck sampling and catch recording procedures

It is important, before the survey begins, to make sure that the equipment and working conditions are such that the sampling can be carried out easily and without risk. Also, the crew must be instructed not to remove any part of the catch before the sampling has been completed.

The following steps pertain to methods for sorting the catch of a fishery research vessel such that the catch composition, both by weight and by number of each species (species group) can be established. The procedure outlined here is from Pauly (1980, adapted from Losse and Dwiponggo, 1977).

Step 1: Remove all sea snakes and other venomous or otherwise dangerous animals. Also remove turtles, and if alive, return these to sea. Record number and kind of animals removed.

Step 2: Remove inorganic debris and plant meterial. Record type of material removed.

Step 3: Remove the larger fish that are readily visible and place them in a box.

Step 4: Wash the remainder of the catch (of small fish) if necessary, and mix with shovels.

Step 5: Put the mixed catch in boxes, while continuing to remove larger fish and putting them into the box mentioned in Step 3. The boxes should be filled simultaneously, not one after the other, and it should be made certain that all boxes contain approximately the same weight of fish.

Step 6: Count the number of boxes with small fish and record.

Step 7: A rule of thumb is to take one box out of every five at random for subsampling. Record number of boxes taken for subsampling as B1, B2, B3, ... etc.

Step 8: The box(es) taken for subsampling is (are) then treated as follows:

- Weigh the total catch in B1 and record.

- Place fish of B1 on a sorting table and sort to species level as far as food fishes and valuable crustaceans (e.g. shrimps) are concerned, and to taxonomic groupings as well-defined as possible (e.g. genus, family, etc.) for other groups (the non-edible fish and miscellaneous crustaceans).

- Repeat procedure, if appropriate, for the other boxes, B2, B3, ... etc.

Step 9: If more than one box were sorted, compute, for each species (or higher taxonomic group) the total weight and number in all sorted boxes.

Step 10: Multiply the numbers and weight of fish and invertebrates by species (or higher taxonomic group) by the ratio of the number of unsorted to sorted boxes.

Step 11: Weigh and count the larger fish mentioned in Steps 3 and 5, by species (very large fish should be weighed inidvidually and measured).

Step 12: Add, when there is an overlap (when the fish of a certain species occurred both in the sorted boxes of small fish and in the large fish box) the weights and numbers obtained in Step 11 to the weights and numbers in Step 10.

Step 13: Step 12 (as well as Step 11, when there is no overlap) provides estimates of total catch, both in weight and number, by species or higher taxonomic groups. Record the total, both in weight and numbers into an appropriate fishing log and convert to catch per unit if fishing time was less or more than an hour. During surveys, this step must be completed after each haul, or every evening at the latest, to preclude loss of information.

Step 14: In addition to catch sampling, identifying and recording, the work of the fishery biologist generally includes among other things:

- Collecting length frequency data
- Collecting miscellaneous biological information on the fish caught, e.g., concerning their weight and maturity
- Collecting and preserving specimens for further studies on-shore
- Collecting oceanographic data.

It is important to sort and sample the catch to the greatest extent possible. The extra cost involved in sorting and sampling all the edible species is usually only a fraction of the total costs of the survey.

13.5 The swept area

From Fig. 13.5.1 it appears that the trawl sweeps a well defined path, the area of which is the length of the path times the width of the trawl, called the "swept area" or the "effective path swept". The swept area, a, can be estimated from:

$$a = D * h * x2, \qquad D = V * t \tag{13.5.1}$$

where V is the velocity of the trawl over the ground when trawling, h is the length of the head-rope (see Fig. 13.5.1), t is the time spent trawling. x2 is that fraction of the head-rope which is equal to the width of the path swept by the trawl, the "wing spread", h * x2 (see Fig. 13.5.1).

For southeast Asian bottom trawls values of x2 from 0.4 (Shindo, 1973) to 0.66 (SCSP, 1978) are reported. Pauly (1980) suggests x2 = 0.5 as the best compromise. In the Caribbean a value of x2 = 0.6 was used by Klima (1976).

For the estimation of biomass we use the catch per unit area (CPUA), as will appear in the next section. It is estimated by dividing the catch by the swept area (in square nautical miles or square kilometers). This estimate thus depends on how accurately the swept area is estimated. The swept area as defined in Fig. 13.5.1 assumes that the bridles have no herding effect, which they are in fact known to exhibit. The wing spread is calculated as the fraction x2 of the head rope length. The wing spread varies with hauling speed, weather conditions, current and warp length and is therefore not well defined. It can be accurately measured by special devices, but an approximate value can be calculated from measurements of the distance between the warps at the gallows and at say, 1 m, towards the net.

When exact positions of the start and the end of the haul are available the distance covered can be estimated in units of nautical miles (nm), by:

$$D = 60 * \sqrt{(Lat1 - Lat2)^2 + (Lon1 - Lon2)^2 * \cos^2(0.5(Lat1 + Lat2))} \tag{13.5.2}$$

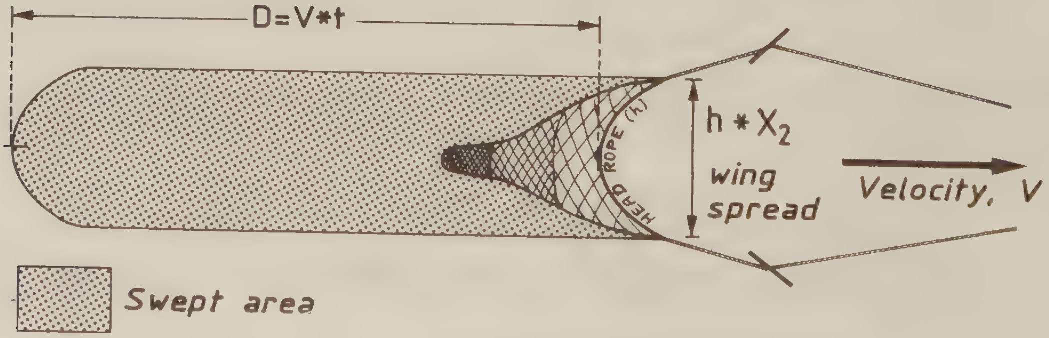


Fig. 13.5.1 Swept area (compare Fig. 4.3.0.1)

where

Lat1 = latitude (degrees) at start of haul
Lat2 = latitude at end of haul
Lon1 = longitude at start of haul
Lon2 = longitude at end of haul

If exact positions are not available, but only the velocity of the vessel and its course together with direction and speed of the current, then the distance covered per hour can be calculated from:

$$D = \sqrt{VS^2 + CS^2 + 2 * VS * CS * \cos(dirV - dirC)} \quad (nm) \qquad (13.5.3)$$

where

VS = velocity of vessel (knots = nm/hr)
CS = velocity of current (knots)
dirV = course of vessel (degrees)
dirC = direction of current (degrees)

Eq. 13.5.3 is an application of "vector addition" (see Fig. 13.5.2)

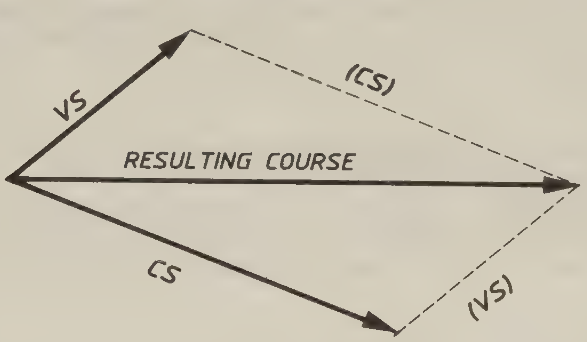


Fig. 13.5.2 Vector addition of vessel and current course and velocity

13.6 Biomass estimation by the swept area method

Let Cw be the catch in weight of a haul. Then Cw/t is the catch in weight per hour, when t is the time spent hauling. Let a be the area swept (cf. Eq. 13.5.1). Then a/t is the area swept per hour, and

$$\frac{Cw/t}{a/t} = \frac{Cw}{a} \quad kg/nm^2 \qquad (13.6.1)$$

becomes the catch in weight per unit of area. Let x1 be the fraction of the biomass in the effective path swept which is actually caught and let $\overline{Cw/a}$ be the mean catch per unit area of all hauls. Then

$$b/a = (\overline{Cw/a})/x1 \quad kg/nm^2 \qquad (13.6.2)$$

is an estimate of the biomass per unit area. Let A nm^2 be the total size of the area under investigation. Then

$$B = \frac{(\overline{Cw/a}) * A}{x1} \qquad (13.6.3)$$

is an estimate of the total biomass in this area.

The proportion of fish in the swept area retained, x1, is difficult to estimate. Underwater television recordings show that the reaction of fish to trawls varies markedly between species. The value of x1 is usually chosen between 0.5 and 1.0. For trawlers used in southeast Asia a value of x1 = 0.5 is commonly used in survey work (Isarankura, 1971, Saeger, Martosubroto and Pauly, 1980). Dickson (1974), on the other hand, suggests x1 = 1. The difference between these two values of x1 is difficult to resolve. Using x1 = 0.5 doubles the estimate of biomass compared to that obtained by using x1 = 1.0.

The duration of the haul is proportional to the distance covered so the duration should have no direct influence on the catch per unit of area. However, the catchability, x1, of different species may vary according to the duration of the haul because some species, when herded by the trawl get tired soon and get captured while other species are able to swim in front of the trawl for a long period and thus avoid being caught. It is therefore important to standardize the duration of the haul so that results from different hauls can be compared. To investigate the dependence of catchability on haul duration, parallel hauls of different duration (e.g half an hour and one hour) could be made.

The total area surveyed can be estimated from a mercator projection by means of a planimeter. Another method is to copy the area on transparent paper which then is cut out and weighed. Then the weight of the paper is compared to the weight of a piece of similar paper which corresponds to a known area. This method is not very precise.

13.7 Precision of the estimate of biomass

Let the biomass estimate in Eq. 13.6.3 be obtained from n hauls, and let Ca(i) be the catch (in weight) per unit of area of haul no. i, i = 1,2,...,n. The estimate of B then becomes:

$$B = \frac{A}{x1} * \frac{1}{n} \sum_{i=1}^{n} Ca(i) = \frac{A}{x1} \overline{Ca} \qquad (13.7.1)$$

and the variance (cf. Eq. 2.3.2):

$$VAR(B) = \left[\frac{A}{x1} \right]^2 \frac{1}{n} \frac{1}{n-1} \sum_{i=1}^{n} [(Ca(i) - \overline{Ca}]^2 \qquad (13.7.2)$$

Thus, a higher precision (a smaller variance) can be obtained by increasing the number of hauls, n (cf. Section 2.7, Fig. 2.7.1). Another way of reducing the variance is to apply stratified sampling (cf. Section 2.8). Suitable stratification may reduce the variance considerably for the same number of hauls and thus improves survey efficiency for the available research vessel time. The distribution of many fish species is determined by depth and bottom type. Therefore stratification based on these factors is widely used. Fig. 13.7.1 shows an example of stratification based on depth.

The areas of the four strata in Fig. 13.7.1 are A1,A2,A3 and A4 and the total area A = A1 + A2 + A3 + A4. Let B(i) be the estimated biomass in stratum no. i calculated by Eq. 13.7.1. Then the estimate of the total biomass of the total area, A becomes:

$$B = B(1) + B(2) + B(3) + B(4) \qquad (13.7.3)$$

and the variance:

$$VAR(B) = VAR(B(1)) + VAR(B(2)) + VAR(B(3)) + VAR(b(4)) \qquad (13.7.4)$$

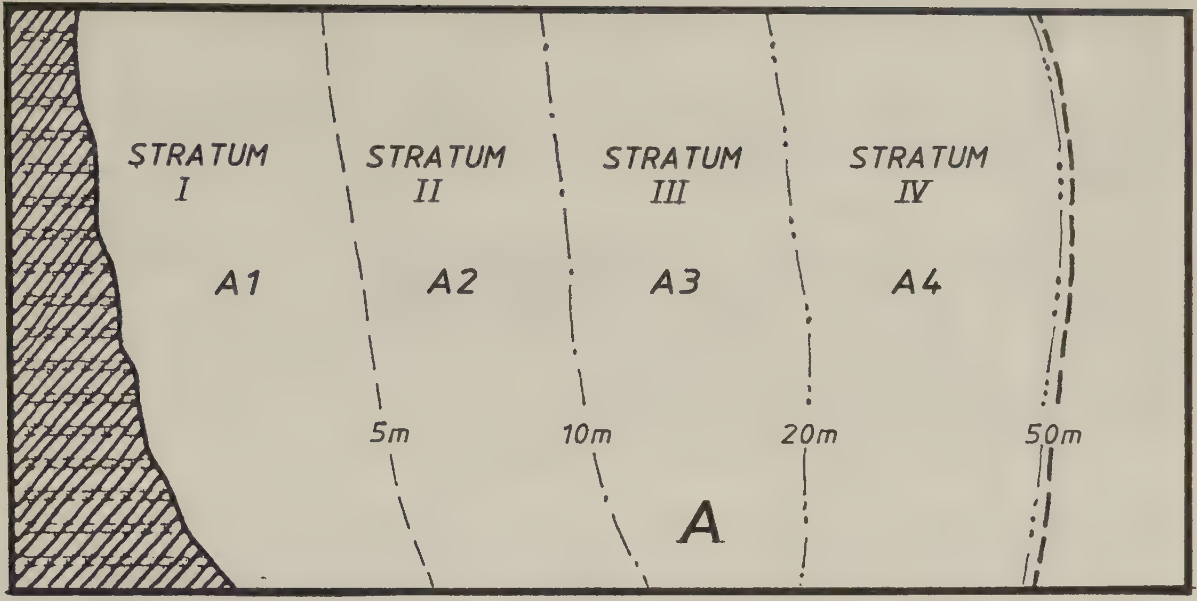


Fig. 13.7.1 Example of stratification

where VAR(B(i)) is calculated by Eq. 13.7.2. If we work in units of biomass per square nm. (cf. Eq. 13.6.2) the parallel to Eq. 13.7.3 becomes:

$$b = (b(1) * A1 + b(2) * A2 + b(3) * A3 + b(4) * A4)/A \qquad (13.7.5)$$

and the parallel to Eq. 13.7.4 is

$$VAR(b) = VAR(b(1)) * (A1/A)^2 + \ldots + VAR(b(4)) * (A4/A)^2 \qquad (13.7.6)$$

In this case a "sample" is a "unit area swept". Thus, here Ai is the size of stratum no. i and A is the total "population" of areas. Eq. 13.7.6 is then seen to be analogous to Eq. 2.8.6 with A as N, Ai as N(j) and the number of unit areas swept in stratum i as n(i). With n(i) small compared to N(i) we can ignore the finite population correction in Eq. 2.8.6 and we get Eq. 13.7.6. Confidence limits of B and b can be calculated as described in Section 2.3.

Stratification may be based on the total catch of a mixture of species or on the catch of a single species. It is often desirable, however, to focus on several species, or groups of species, each with its own type of distribution. In such cases a stratification procedure must be decided on for each species, or for each group of species with similar distribution patterns within the area surveyed.

For a more detailed description of the statistical aspects of a trawl survey see Fogarty (1985) and Gavaris and Smith (1987). The latter discuss the effect of stratification by comparing the precision obtained from stratified sampling to that of simple random sampling, using the method of Sukhatme and Sukhatme (1970).

The above formulas all assume that the number caught per trawl hour is a normally distributed random variable. A better description of survey data is often obtained by assuming the so-called "delta-distribution" (Pennington, 1983). This distribution takes into account that survey data often contain a large proportion of zero observations and that the non-zero observations form a log-normal distribution rather than a normal distribution.

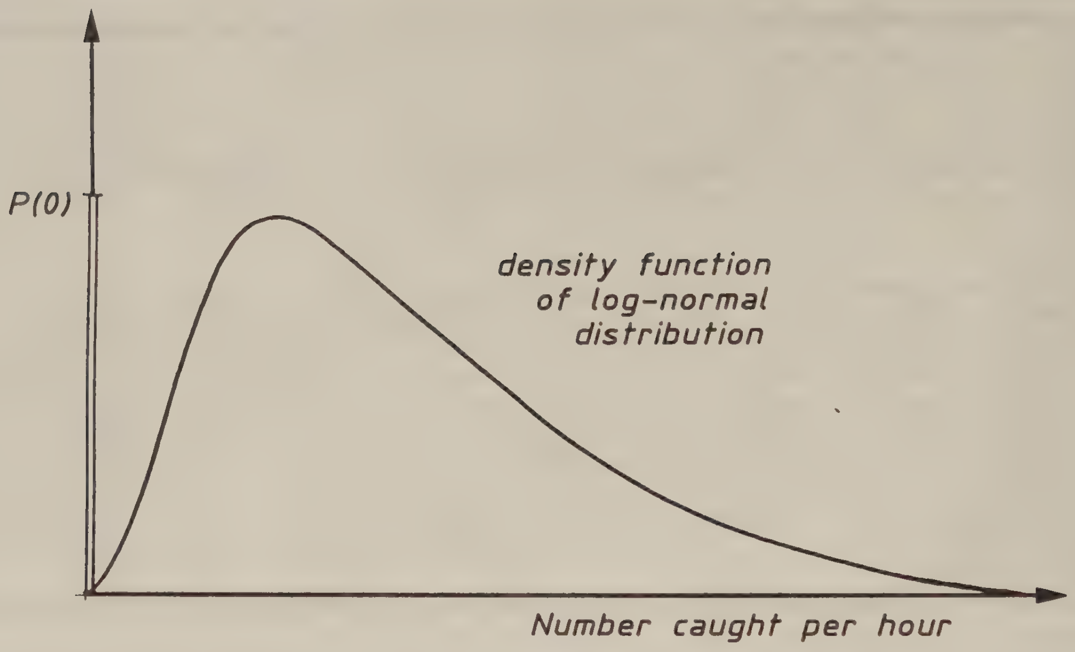


Fig. 13.7.2 Delta-function

A log-normal distribution is asymmetric and skewed to the right hand side. Fig. 13.7.2 shows a schematic delta-function. Note that the point zero has a positive probability. In his paper, Pennington gives the expression for the unbiased estimators of the parameters of the delta-function.

13.8 Estimation of maximum sustainable yield

Gullands formula for virgin stocks (cf. Section 9.3):

$$MSY = 0.5 * M * Bv$$

(M: coefficient of natural mortality; Bv: virgin stock) or Cadima's formula for exploited stocks (cf. Section 9.4):

$$MSY = 0.5 * 2 * B \quad \text{or} \quad MSY = 0.5 * (Y + M * B)$$

have often been applied when estimating the maximum sustainable yield (MSY) per year from the total biomass as estimated from exploratory surveys (see e.g. Saville, 1977, or Troadec, 1980). It may be better, however, to apply the modified Fox model (Garcia, Sparre and Csirke, 1987) described in Section 9.5 using Eq. 9.5.10 or Eq. 9.5 11.

Anyway, these methods are crude ones and should be relied upon to give only the order of magnitude (1000 tonnes, 10,000 tonnes, etc.) of the fish stock. Yet, if nothing is known in advance, even such an estimate of the order of magnitude may be valuable.

One problem is that M is usually not known before the surveys are carried out. When the stock is virgin (unexploited) an estimate of M can be obtained by analyzing the catch curves obtained during the exploratory survey (cf. Section 4.4) because when the stock is unexploited, then F = 0 and thus Z = M. Pauly's empirical formula, Eq. 4.7.2.1 (Pauly, 1980b) can be applied when growth parameters are available.

(See **Exercise(s)** in Chapter 17).

14 LFSA PACKAGE OF MICROCOMPUTER PROGRAMS

The LFSA (**L**ength based **F**ish **S**tock **A**ssessment) package of microcomputer programs (Sparre, 1987), written in the language BASIC, is designed to match the present manual. Most of the programs follow exactly the procedure explained in this manual.

The LFSA package was originally developed for Apple II computers as indicated by the title of its manual (Sparre, 1987). Later, these programs were converted for use on IBM compatible computers. The IBM version has been updated and considerably expanded whereas the Apple II version has not. Future developments will probably lead to merging the LFSA package with the COMPLEAT ELEFAN package developed by ICLARM (Gayanilo, Soriano and Pauly, 1988), but for the time being the two are separate, although data can easily be transferred from LFSA to ELEFAN and vice versa.

The LFSA package runs on Apple II and IBM (PC, XT or AT) microcomputers or their compatibles with minimum memory, a monochrome screen and an 80 columns printer. The graphics of the IBM version is based on the IBM Color-Graphics Monitor Adapter (High resolution graphics, 640 X 200).

Not all the methods given in this manual have been computerized. The emphasis has been placed on incorporating the time consuming operations, such as the Bhattacharya analysis. A high degree of user-friendliness has been attempted. The user is neither assumed to have knowledge of the BASIC language nor of the operation system used (PRODOS for Apple II and MS-DOS for IBM). The programs are executed via self-explanatory screen dialogues. The prompts often take the form of a "menu", from which the user selects a "course". In addition to the programs dealing with the methods described in this manual, the LFSA package also contains a number of utility programs for data handling, and initial data processing.

14.1 Length frequency (LF) programs

Fig. 14.1 shows a flowchart of the most important part of the LFSA package, namely the subpackage for length frequency analysis.

LFINPUT: LF data entry/edit program.
The program is used to enter/edit time series of length frequencies and store these on magnetic disk (the type of data required for e.g. modal progression analysis, cf. Section 3.4.2). LFINPUT automatically sorts the length frequency samples in chronological order.

LFINF: LF additional data program.
This program is an optional continuation of LFINPUT, which allows the attachment of a variety of additional data to each length frequency sample. Such additional data may be effort, total catch, position, depth, gear type, bottom type, ... etc.

LFTABLE: LF table program.
Produces tables of time series of length frequencies formatted according to the user's choice.

LFBAR: LF bar diagram program.
Produces low resolution histograms, using only typewriter characters (mainly for use by Apple II version).

LFGRAPH: LF bar diagram program.
The high resolution counterpart to LFBAR (only in IBM version).

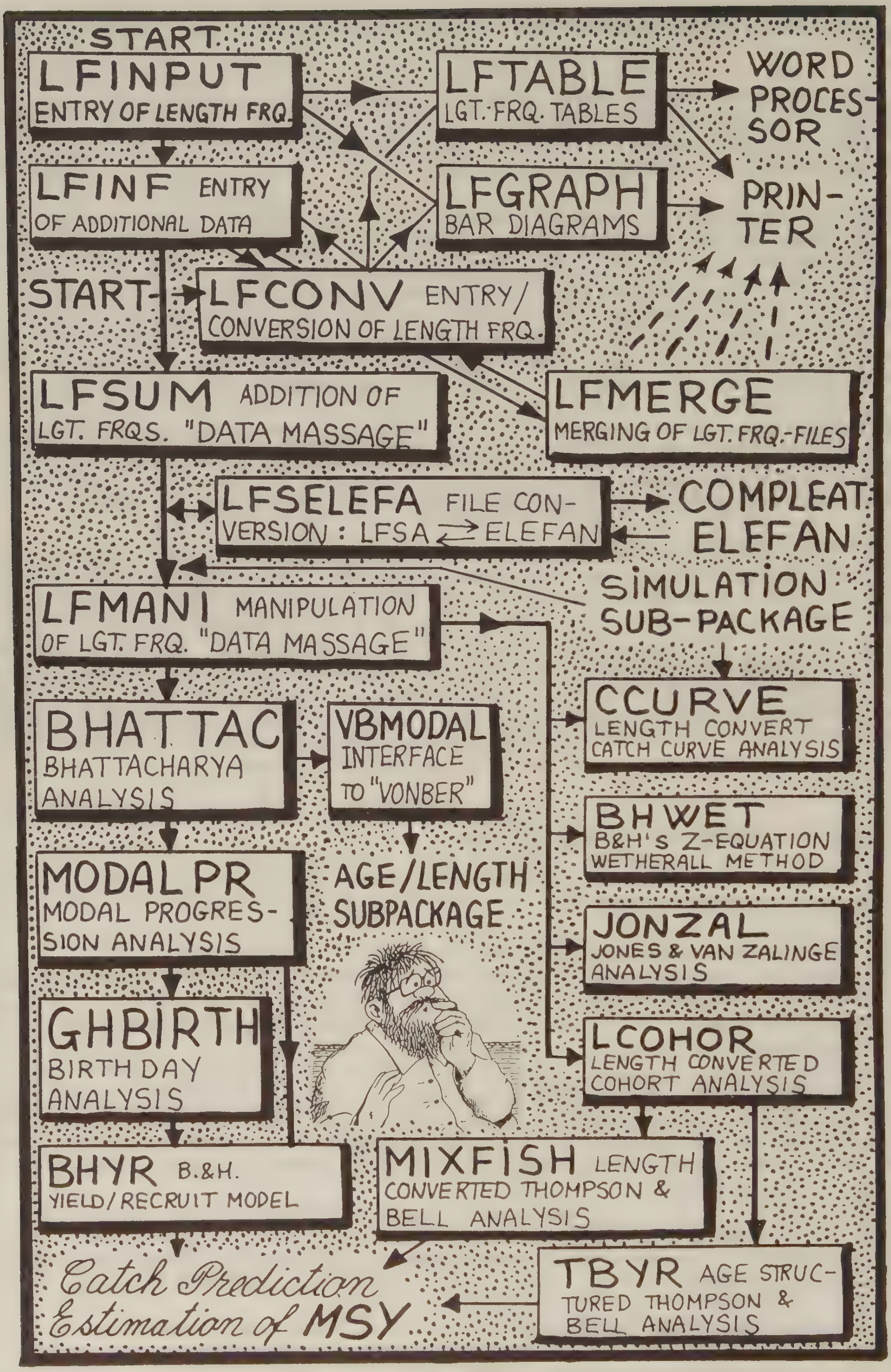

START
LFINPUT
ENTRY OF LENGTH FRQ.
LFTABLE
LGT. FRQ. TABLES
WORD PROCESSOR
LFINF ENTRY OF ADDITIONAL DATA
LFGRAPH
BAR DIAGRAMS
PRINTER
START
LFCONV ENTRY/ CONVERSION OF LENGTH FRQ.
LFSUM ADDITION OF LGT. FRQS. "DATA MASSAGE"
LFMERGE
MERGING OF LGT. FRQ.-FILES
LFSELEFA FILE CONVERSION : LFSA ⇄ ELEFAN
COMPLEAT ELEFAN
SIMULATION SUB-PACKAGE
LFMANI MANIPULATION OF LGT. FRQ. "DATA MASSAGE"
BHATTAC
BHATTACHARYA ANALYSIS
VBMODAL
INTERFACE TO "VONBER"
CCURVE
LENGTH CONVERT CATCH CURVE ANALYSIS
BHWET
B&H'S Z-EQUATION WETHERALL METHOD
MODALPR
MODAL PROGRESSION ANALYSIS
AGE/LENGTH SUBPACKAGE
JONZAL
JONES & VAN ZALINGE ANALYSIS
GHBIRTH
BIRTH DAY ANALYSIS
LCOHOR
LENGTH CONVERTED COHORT ANALYSIS
BHYR B.&H. YIELD/RECRUIT MODEL
MIXFISH LENGTH CONVERTED THOMPSON & BELL ANALYSIS
TBYR AGE STRUCTURED THOMPSON & BELL ANALYSIS
Catch Prediction
Estimation of MSY

Fig. 14.1.1 Flowchart for the major programs of the LFSA subpackage for length frequency analysis (not all programs are included)

LFSUM: LF data summing program.
This program contains a number of options for raising and summing length frequency samples (cf. Chapter 7).

LFCONV: LF data conversion program.
In practice it may happen that data from different sources are incompatible. E.g. some samples may be in units of fork length and some samples in units of total length (cf. Section 1.5). Although the formulas for conversion of one measurement into another appear quite simple, it is not a simple thing to convert grouped data. LFCONV can solve this problem and some similar problems. LFCONV can also be used to convert commercial size categories (e.g number of shrimp tails per pound) into length frequencies. LFCONV is an extended version of LFINPUT.

LFMANI: LF data manipulation program
This program offers a number of options, for example to calculate the weight of a sample based on the length/weight relationship (cf. Section 2.6). It may also be used to convert from, say, 1 cm length classes into 2 cm length classes, to convert into relative frequencies, to correct for gear selection (cf. Section 6.5) etc.

LFMERGE: LF file merging program.
This program is used to merge different files of time series. It merges files so that the resulting file has its samples in chronological order.

LFSELEFA: File conversion LFSA/ELEFAN.
Converts length frequency files of LFSA-format into ELEFAN-format-files and ELEFAN-files into LFSA-files. "ELEFAN-files" means files which can be used as input to the "COMPLEAT ELEFAN" package of microcomputer programs (Gayanilo, Soriano and Pauly, 1988). This program allows you to change from one package to the other without having to reenter the input data.

BHATTAC: Bhattacharya analysis program.
Performs the Bhattacharya analysis as described in this manual (Section 3.4.1). When executed by paper and pencil assisted by pocket calculator this analysis is quite time consuming (cf. Exercise 3.4.1). With BHATTAC it would take only a few minutes to solve it. The graphs are displayed on the screen and can optionally be converted into a hard copy.

BHGRAPH: Display of results of BHATTAC in high resolution graphics (only in IBM version). The graphs produced by BHGRAPH resemble Fig. 3.2.2.2.

MODALGR: Bhattacharya graph and results.
Produces a graph similar to Fig. 3.4.2.1 and a corresponding table (cf. Table 3.4.2.1). Input to MODALGR is the output from program BHATTAC.

MODALPR: Modal progression analysis.
The analysis described in the first part of Section 3.4.2, i.e. the estimation of growth parameters by the Gulland and Holt plot (cf. Table 3.4.2.3 and Fig. 3.4.2.2) based on the results of the Bhattacharya analysis. Input to MODALGR is the output from program BHATTAC.

VBMODAL: Interface to program VONBER.
This program is an alternative to MODALPR. The user is asked to group the components found with BHATTAC into cohorts (similar to modal progression analysis). The result of this exercise is a series of pairs:

(L(1),t(1),(L(2),t(2)),.....,(L(n),t(n))
where L(i) = length of fish no. i.
t(i) = arbitrary age of fish no. i.

These data are in turn used as input to the program "VONBER", which estimates the growth parameters.

GHBIRTH: Birthday analysis.
The analysis described in the second part of Section 3.4.2, i.e. the estimation of approximate birthday by the van Bertalanffy plot (cf. Table 3.4.2.4 and Figs. 3.4.2.3 and 3.4.2.4). Input to MODALGR is the output from program MODALPR.

CCURVE: Length converted catch curve analysis.
The analysis described in Sections 4.4.5 and 6.1.2, i.e. estimation of Z and estimation of the gear selection ogive from the catch curve.

JONZAL: Jones and van Zalinge method.
Estimation of Z as described in Section 4.4.6.

BHZWET: Beverton and Holt's Z-equation and the modified Wetherall et al. analysis.
Estimation of Z as described in Section 4.5.2 and estimation of L_∞ and Z/K as described in Section 4.5.4.

LCOHOR: Jones' length based cohort analysis.
Estimation of stock size and fishing mortality as described in Section 5.4.

TBYR: Age based Thompson and Bell model yield calculations.
This program uses a special version of the Thompson and Bell yield and stock prediction model for the single stock single fishery situation (cf. Section 5.3). TBYR takes its starting point in the stock numbers by length group calculated by LCOHOR and converts them into age groups. The following calculations are the traditional ones as descrbied in Section 5.3. Because the conversion from length groups to age groups (whole years) is problematic for short lived species, this program should be used only for long lived species (5 years or more).

MIXFISH: Length based Thompson and Bell model with option for analysis of a mixed fishery (only in IBM version).
This program is based on the theory explained in Sections 5.5 and 10.4.1). It does essentially the same as TBYR, but without the conversion of lengths into age groups. It may be used for long lived as well as short lived species. Although designed for analysis of a mixed fishery, MIXFISH contains the single species case as an option.

SIMULL: Monte Carlo simulation of length frequency samples (only in IBM version).
The program SIMULL simulates length frequency samples by simulating the life history of each individual fish, using Monte Carlo techniques.

SIMMIG: Monte Carlo simulation of length frequency samples accounting for migration (only in IBM version).
The program SIMULL simulates length frequency samples the same way as SIMULL. It does the simulation for several areas in parallel and simulates the movements of fish between the areas.

14.2 Age/length (AL) analysis: estimation of growth parameters from age/length data

VBINPUT: Age/length (AL) data entry program.
Resembles program LFINPUT.

VONBER: Estimation of von Bertalanffy parameters.
This program estimates the growth parameters in the ordinary von Bertalanffy growth equation (Eq. 3.1.0.1) by non-linear regression analysis as described in Section 3.1. Because the computation procedure is based on iteration techniques requiring a very large number of calculations this method has not been dealt with as an example in this manual.

The input to VONBER is a series of pairs:

(L(1),t(1)),(L(2),t(2)),.....,(L(n),t(n))

where

L(i) = length of fish no. i.
t(i) = age of fish no. i.

Age data may be obtained from reading hard parts (otoliths, scales, ..etc.).

The program VONBER determines the parameters, L_∞, K and t_o so that the sum of squares of deviations:

$$\sum_{i=1}^{n} [L(i) - L_\infty\{1 - \exp(-K(t(i)-t_o))\}]^2$$

is minimized.

VBRESU: Results from VONBER.

VBGRAPH: Graph of results from VONBER.

VBMERGE: Age/length file merging program.

14.3 Various other programs

BHYR: Beverton and Holt's yield per recruit analysis.
The estimation of Yield/Recruit and Biomass/Recruit as described in Chapter 8.

PAULYM: Pauly's M-formula (only in IBM version).
Calculates M, using Pauly's formula (Eq. 4.7.2.1).

REGRES: Ordinary regression analysis.
Performs the type of analysis described in Sections 2.4 to 2.6. REGRES contains a number of options for linearization.

REGRAPH: Graph for ordinary regression analysis.

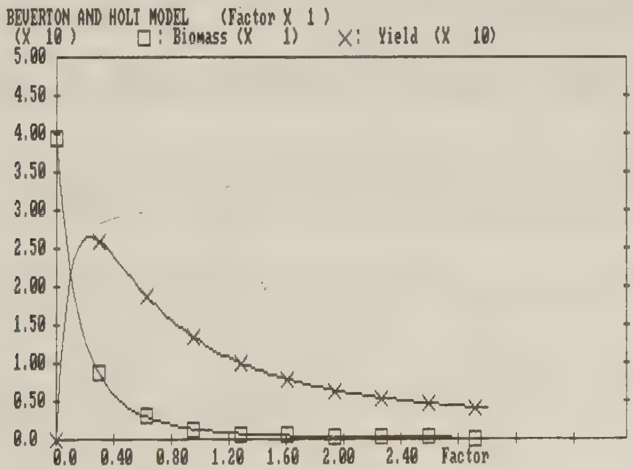


Fig. 14.1.2 Example of a print out of the BHYR program

15 REFERENCES

Alagaraja, K., 1984. Simple methods for estimation of parameters for assessing exploited fish stocks. *Indian J.Fish.*, 31:177-208

Allan, K.R. and G.P. Kirkwood, 1988. Marine mammals. *In* Fish population dynamics: the implications for management, edited by J.A. Gulland, Chichester, John Wiley and Sons Ltd., pp. 251-69

Allen, G.R., 1985. FAO species catalogue. Vol. 6. Snappers of the world. An annotated and illustrated catalogue of lutjanid species known to date. *FAO Fish.Synop.*, (125)Vol.6:208 p.

Alverson, D.L. (ed.), 1971. Manual of methods for fisheries resource survey and appraisal. Part 1. Survey and charting of fisheries resources. *FAO Fish.Tech.Pap.*, (102):80 p.

Alverson, D.L. and W.T. Pereyra, 1969. Demersal fish explorations in the northeastern Pacific Ocean - An evaluation of exploratory fishing methods and analytical approaches to stock size and yield forecasts. *J.Fish.Res.Board.Can.*, 26:1985-2001

Amaratunga, T., 1983. The role of cephalopods in the marine ecosystem. *FAO Fish.Tech.Pap./FAO Doc.Tech.Pêches/FAO Doc.Téc.Pesca*, (231):379-415

Andersen, K.P. and E. Ursin, 1977. A multispecies extension to the Beverton and Holt theory of fishing, with accounts of phosphorus circulation and primary production. *Medd.Dan.Fisk.-og Havunders.(Ny Ser.)*, 7:319-435

Australian Journal of Marine and Freshwater Research, 1987. Prawn ecology and biology. *Aust.J.Mar.Freshwat.Res.*, 38(1):190 p.

Bagenal, T.B., (ed.), 1974. Ageing of fish. Proceedings of an international symposium on the ageing of fish held at the Univ. of Reading, England 19-20 July 1973. Old Woking, Surrey, England, Unwin Bros. Ltd., 234 p.

Bakhayokho, M., 1983. Biology of the cuttlefish *Sepia officinalis hierredda* off the Senegal coast. *FAO Fish.Tech.Pap./FAO Doc.Tech.Pêches/FAO Doc.Téc.Pesca.*, (231):204-63

Bakun, A., *et al.*, 1982. Ocean sciences in relation to living resources. *Can.J.Fish.Aquat.Sci.*, 39:1059-70

Baranov, F.I., 1914. The capture of fish by gillnets. *Mater.Poznaniyu Russ.Rybolov.* 3(6):56-99 (Partially translated from Russian by W.E. Ricker)

Baranov, F.I., 1918. On the question of the biological basis of fisheries. *Nauchn.Issled.Ikhtiol.Inst.Izv.*, 1:81-128 (in Russian)

Baranov, F.I., 1926. On the question of the dynamics of the fishing industry. *Nauchn.Byull.Rybn.Khoz.*, 8(1925):7-11 (in Russian)

Baranov, F.I., 1948. Theory and assessment of fishing gear. Moscow, Pishchepromizdat. English version issued as Theory of fishing with gill nets. Transl. from Russian by Ontario Department of Land and Forests, Toronto, 45 p.

Bard, F.X., 1984. Le listao de l'Atlantique. 3. Resultats des campagnes de marquages effectuées de 1980 a 1982. *Pêche Marit.*, 1275:319-24

Bard, F.X., P. Cayre and T. Diouf, 1988. Les migrations. *FAO Doc.Tech.Pêches*, (292):111-156

Bayliff, W.H. and K.N. Holland, 1986. Materials and methods for tagging tuna and billfishes, recovering the tags and handling the recapture data. *FAO Fish.Tech.Pap.*, (279):36 p.

Beamish, R.J. and G.A. McFarlane, 1983. The forgotten requirement for age validation in fisheries biology. *Trans.Am.Fish.Soc.*, 112(6):735-43

Beddington, J.R. and J.G. Cooke, 1983. The potential yield of fish stocks. *FAO Fish.Tech.Pap.*, (242):47 p.

Berkeley, S.A. and E.D. Houde, 1980. Swordfish, *Xiphias gladius*, dynamics in the Straits of Florida. ICES C.M. 1980/H:59:11 p. (mimeo)

Berkeley, S.A., 1983. Atlantic swordfish stock structure data and suggestions for its interpretation. *Collect.Vol.Sci.Pap.ICCAT/Recl.Doc.-Sci.CICTA/Colecc.Doc.Cient.CICAA*, 18(3):839-45

Bertalanffy, L. von, 1934. Untersuchungen über die Gesetzlichkeiten des Wachstums. 1. Allgemeine Grundlagen der Theorie. *Roux'Arch.Entwicklungsmech.Org.*, 131:613-53

Beverton, R.J.H., 1963. Maturation, growth and mortality of clupeid and engraulid stocks in relation to fishing. *Rapp.P.-V.Réun.CIEM*, 154:44-67

Beverton, R.J.H. and S.J. Holt, 1956. A review of methods for estimating mortality rates in exploited fish populations, with special reference to sources of bias in catch sampling. *Rapp.P.-V.Réun.CIEM*, 140:67-83

Beverton, R.J.H. and S.J. Holt, 1957. On the dynamics of exploited fish populations. *Fish.Invest.Minist.Agric.Fish.Food G.B.(2 Sea Fish.)*, 19:533 p.

Beverton, R.J.H. and S.J. Holt, 1959. A review of the lifespans and mortality rates of fish in nature, and their relation to growth and other physiological characteristics. *In* CIBA Foundation, colloquia on ageing. Vol. 5. The lifespan of animals, edited by G.E.W. Wolstenholme and M. O'Connor. London, Churchill, Vol.5:142-80

Beverton, R.J.H. and S.J. Holt, 1966. Manual of methods for fish stock assessment. Part 2. Tables of yield functions. Manuel sur les méthodes d'évaluation des stocks ichtyologiques. Partie 2. Tables de fonctions de rendement. Manual de métodos para la evaluación de los stocks de peces. Parte 2. Tablas de funciones de rendimiento. *FAO Fish.Tech.Pap./FAO Doc.Tech.Pêches/FAO Doc.Téc.Pesca*, (38)Rev.1:67 p.

Beyer, J. and P. Sparre, 1983. Modelling exploited fish stocks. *In* Application of ecological modelling in environmental management. Part A, edited by S.E. Jørgensen. Amsterdam, Elsevier Scientific Publishing Co., pp. 485-582

Bhattacharya, C.G., 1967. A simple method of resolution of a distribution into Gaussian components. *Biometrics*, 23:115-35

Blaber, S.J.M. and T.G. Blaber, 1980. Factors affecting the distribution of juvenile estuarine and inshore fish. *J.Fish Biol.*, 17:143-62

Booke, H.E., 1981. The conundrum of the stock concept - are nature and nurture definable in fishery science? *Can.J.Fish.Aquat.Sci.*, 38:1479-80

Boonyubol, M. and V. Hongskul, 1978. Demersal resources and exploitation in the Gulf of Thailand, 1960-1975. In Report on the Workshop on the demersal resources of the Sunda Shelf, Penang, Malaysia. Manila, South China Sea Fisheries Development and Coordinating programme, SCS/GEN/77/13:56-70

Brander, K., 1988. Multispecies fisheries of the Irish Sea. In Fish population dynamics: the implications for management, edited by J.A. Gulland, Chichester, John Wiley and Sons Ltd., pp. 303-28

Brody, S. 1927. Growth rates. Bull.Univ.Miss.Agric.Exp.Sta., (97)

Brothers, E.B., 1980. Age and growth studies on tropical fishes. In Stock assessment for tropical small-scale fisheries. Proceedings of an International Workshop held 19-21 September at the University of Rhode Island, Kingston, R.I., edited by S.B. Saila and P. Roedel International center for Marine Resource Development, University of Rhode Island, pp. 119-36

Butler, M.J.A. et al., 1986. Marine resource mapping: an introductory manual. FAO Fish.Tech.Pap., (274):256 p.

Caddy, J.F., 1980. Surplus production models. In Selected lectures from the CIDA/FAO/CECAF seminar on fishery resource evaluation. Casablanca, Morocco, 6-24 March 1978: Rome, FAO Canada Funds-in-Trust, FAO/-TF/INT 180(c) Suppl.: 29-55. Issued also in French

Caddy, J.F., 1982. Some considerations relevant to the definition of shared stocks and their allocation between adjacent economic zones. FAO Fish.Circ., (749):44 p. Issued also in Spanish

Caddy, J.F. (ed.), 1983. Advances in assessment of world cephalopod resources. Progrès réalisés dans l'évaluation des ressources mondiales de céphalopodes. Progresos realizados en la evaluación mundial de cefalópodos. FAO Fish.Tech.Pap./FAO Doc.Tech.Pêche/FAO Doc.Téc.Pesca, (231):452 p.

Caddy, J.F., 1983a. The cephalopods: factors relevant to their population dynamics and to the assessment and management of stocks. FAO Fish.-Tech.Pap./FAO Doc.Tech.Pêche/FAO Doc.Téc.Pesca, (231): 416-52

Caddy, J.F., 1987. Size-frequency analysis for crustacea: moult increment and frequency models for stock assessment. Kuwait Bull.Mar.Sci., (9):43-61

Caddy, J.F., 1987a. Types of resource sharing relevant to Lesser Antilles stocks. FAO Fish.Rep., (383):115-24

Caddy, J.F., 1989. Marine invertebrate fisheries: their assessment and management. New York, John Wiley and Son Inc., Interscience, 752 p.

Caddy, J.F. and J. Csirke, 1983. Approximations to sustainable yield for exploited and unexploited stocks. Océanogr.Trop., 18(1):3-15

Caddy, J.F. and S. Garcia, 1986. Fisheries thematic mapping - a prerequisite for intelligent management and development of fisheries. Océanogr.Trop., 21(1):31-52

Carey, F.G. and R.J. Olson, 1982. Sonic tracking experiments with tunas. Collect.Vol.Sci.Pap.ICCAT/Recl.Doc.Sci.CICTA/Colecc.Doc.Cient.CICAA, 17:458-66

Cassie, R.M., 1954. Some uses of probability paper in the analysis of size frequency distributions. Aust.J.Mar.Freshwat.Res., 5:513-22

CECAF, 1978. Report of the Ad hoc working group on hake (Merluccius merluccius, Merluccius senegalensis, Merluccius cadenati) in the northern zone of CECAF held at the Instituto Español de Oceanografia, Santa Cruz de Tenerife, Canary Islands, Spain, 5-9 June 1978. CECAF/ECAF Ser., (48/9): 93 p. Issued also in French

Chapman, D.G. and D.S. Robson, 1960. The analysis of a catch curve. Biometrics, 16(3):354-68

Chikuni, S., 1983. Cephalopod resources in the Indo-Pacific region. FAO Fish.Tech.Pap./FAO Doc.Tech.Pêches/FAO Doc.Téc.Pesca, (231):264-305

CIDA/FAO/CECAF, 1980. Selected lectures from the CIDA/FAO/CECAF seminar on fishery resource evaluation. Casablanca, Morocco, 6-24 March 1978. Rome, FAO, Canada Funds-in-Trust, FAO/TF/INT 180(c) Suppl.:166 p.

Clark, W.G., 1981. Restricted least-squares estimates of age composition from length composition. Can.J.Fish.Aquat.Sci., 38:297-307

Cloern, J.E. and F.H. Nichols, 1978. A von Bertalanffy growth model with a seasonally varying coefficient. J.Fish.Res.Board Can., 35:1479-82

Cochran, W.G., 1977. Sampling techniques. New York, John Wiley and Sons, Inc., 428 p. 3rd ed.

Colette, B.B. and C.E. Nauen, 1983. FAO species catalogue. Vol. 2. Scrombrids of the world. An annotated and illustrated catalogue of tunas, mackerels, bonitos and related species known to date. FAO Fish.Synop., (125)Vol.2:137 p.

Compagno, L.J.V., 1984. FAO species catalogue. Vol.4. Sharks of the world. An annotated and illustrated catalogue of shark species known to date. Part 1. Hexanchiformes to Lamniformes. FAO Fish.Synop., (125)Vol.4,Pt.1:249 p.

Compagno, L.J.V., 1984a. FAO species catalogue. Vol.4. Sharks of the world. An annotated and illustrated catalogue of shark species known to date. Part 2. Carcharhiniformes. FAO Fish.Synop., (125)Vol.4,Pt.2: 251-655

Csirke, J., 1980. Recruitment in the Peruvian anchovy and its dependence on the adult population. Rapp.P.-V.Réun.CIEM, 177:307-13

Csirke, J., 1980a. Introducción a la dinámica de problaciones de peces. FAO Doc.Téc.Pesca, (192):82 p.

Csirke, J., 1988. Small shoaling pelagic fish stocks. In Fish population dynamics: the implications for management, edited by J.A. Gulland, Chichester, John Wiley and Sons Ltd., pp. 271-302

Cushing, D.H., 1968. Fisheries biology: a study in population dynamics. Madison, University of Wisconsin Press, 200 p.

Cushing, D.H., 1988. The study of stock and recruitment. In Fish population dynamics: the implications for management, edited by J.A. Gulland, Chichester, John Wiley and Sons Ltd., pp. 105-28

Dawson, W.A., 1986. Changes in western mackerel (Scomber scombrus) spawning stock composition during the spawning season. J.Mar.Biol.Assoc.U.K., 66:367-83

Dayaratne, P. and J. Gjøsaeter, 1986. Age and growth of four Sardinella species from Sri Lanka. Fish.Res., 4:1-33

Deriso, R.B., 1980. Harvesting strategies and parameter estimation for an age-structured model. Can.J.Fish.Aquat.Sci., 37:268-82

Dickson, W., 1974. A review of the efficiency of bottom trawls. Bergen, Norway, Institute of Fisheries Technology and Research, 44 p.

Doubleday, W.G., 1980. Survey methods for resource evaluation. In Selected lectures from the CIDA/FAO/CECAF seminar on fishery resource evaluation. Casablanca, Morocco, 6-24 March 1978. Rome, FAO, Canada Funds-in-Trust, FAO/TF/INT 180(c) Suppl:65-71

Doubleday, W.G. (ed.), 1981. Manual on groundfish surveys in the Northwest Atlantic. NAFO Sci.Counc.Stud., (2):7-55

Dwiponggo, A., 1979. Review of the demersal resources and fisheries in the Java Sea. IPFC:RRD/II/79/Inf.12. Paper presented at the SCORRAD Meeting, 1979, Hongkong

Edser, T., 1908. Note on the number of plaice at each length in certain samples from the southern part of the North Sea, 1906. J.R.Stat.Soc., 71:686-90

Edwards, R.R.C., 1985. Growth rates of Lutjanidea (snappers) in tropical Australian waters. J.Fish Biol., 26:1-4

Elliott, J.M., 1971. Some methods for the statistical analysis of samples of benthic invertebrates. Sci.Publ.Freshwat.Biol.Assoc.,Ambleside,-U.K., (25):144 p.

Eltink, A.T.G.W., 1987. Changes in age-size distribution and sex ratio during spawning and migration of western mackerel (Scomber scombrus L.). J.Cons.CIEM, 44:10-22

Ehrhardt, N.M. et al., 1983. On the fishery and biology of the giant squid Dosidicus gigas in the Gulf of California, Mexico. FAO Fish.Tech.-Pap./FAO Doc.Tech.Pêches/FAO Doc.Téc.Pesca, (231):306-39 pp

FAO, 1978. Models for fish stock assessment. Partial translation of the annex to the report of the second FAO/CNEXO Training Centre on methods for fish stock assessment, Brest (France), 26 July-27 August 1976. Original French edition 164 p. FAO Fish.Circ., (701): 122 p.

FAO, 1978 a. Some scientific problems of multispecies fisheries. Report of the expert consultation on management of multispecies fisheries. FAO Fish.Tech.Pap., (181):42 p.

FAO, 1978 b. FAO catalogue of fishing gear designs. Catalogue FAO de plans d'engins de pêche. Catalogo de la FAO de planos de aparajos de pesca. Farnham, Surrey, Fishing News Books for FAO, 159 p.

FAO/IOP, 1979. Report of the FAO/IOP Workshop on the fishery resources of the western Indian Ocean south of the equator. Rome, FAO/UNDP, Indian Ocean Programme, IOFC/DEV/79/45:102 p.

FAO/UNDP, 1975. FAO regional fishery survey and development project, Doha (Qatar). Report of the Ad hoc working group on survey technique and strategy. Rome, FAO, FI:DP/REM/71/278/1: 45 p.

Fischer, W., (ed.), 1978. FAO species identification sheets for fishery purposes. Western Central Atlantic. (fishing area 31). Rome FAO, 7 vols: pag.var

Fischer, W. and G. Bianchi (eds), 1984. FAO species identification sheets for fishery purposes. Western Indian Ocean (fishing area 51). Prepared and printed with the support of the Danish International Development Agency (DANIDA). Rome, FAO, 6 vols. pag.var.

Fischer, W., G. Bianchi and W.B. Scott (eds), 1981. FAO species identification sheets for fishery purposes. Eastern Central Atlantic; fishing areas 34, 47(in part). Canada Funds-in-Trust. Ottawa, Department of Fisheries and Oceans Canada, by arrangement with the Food and Agriculture Organization of the United Nations, 7 vols: pag.var.

Fischer, W. and J.C. Hureau (eds), 1985. FAO species identification sheets for fishery purposes. Southern Ocean (Fishing areas 48, 58 and 88) (CCAMLR Convention Area). Prepared and published with the support of the Commission for the Conservation of Antarctic Marine Living Resources. Rome, FAO, 2 vols:470 p. Issued also in French and Spanish.

Fischer, W., M. Schneider and M.-L. Bauchot, (rédacteurs), 1987. Fiches FAO d'identification des espèces pour les besoins de la pêche. (Révision 1). Méditerranée et Mer Noire. Zone de pêche 37. Publication préparée par la FAO, résultat d'un accord entre la FAO et la Commission des Communautés Européennes (Project GCP/INT/422/EEC) financée conjointement par ces deux organisations. Rome, FAO, 2 vols:1532 p.

Fischer, W. and P.J.P. Whitehead (eds), 1974. FAO species identification sheets for fishery purposes. Eastern Indian Ocean (fishing area 57) and Western Central Pacific (fishing area 71). Rome, FAO, 4 vols:pag.var.

Flowers, J.M., 1978. A data processing and basic analysis system for demersal fisheries surveys. Regional fishery survey and development project. Rome, FAO, FI:DP/REM/71/278/4: 151 p.

Fogarty, M.J., 1985. Statistical considerations in the design of trawl surveys. *FAO Fish.Circ.*, (786):21 p.

Fonteneau, A., 1982. Eléments pour l'aménagement des pêcheries d'albacore (*Thunnus albacares*) de l'Atlantique. *Collect.Vol.Sci.Pap.ICCAT/Recl.Doc.Sci.CICTA/Colecc.Doc.Cient.CICAA*, 17:79-163

Fonteneau, A. and J. Marcille, (eds), 1988. Ressources, pêche et biologie des thonidés tropicaux de l'Atlantique centre est. *FAO Doc.Tech.Pêches*, (292):391 p.

Ford, E., 1933. An account of the herring investigations conducted at Plymouth during the years from 1924 to 1933. *J.Mar.Biol.Assoc.U.K.*, 19:305-84

Forsbergh, E.D., 1980. Synopsis of biological data on the skipjack tuna, *Katsuwonus pelamis* (Linnaeus, 1758), in the Pacific Ocean. *Spec.Rep.I-ATTC*, (2):295-360

Fox, W.W. Jr., 1970. An exponential surplus-yield model for optimizing exploited fish populations. *Trans.Am.Fish.Soc.*, 99:80-8

Francis, R.C., 1974. Relationship of fishing mortality to natural mortality at the level of maximum sustainable yield under the logistic stock production model. *J.Fish.Res.Board Can.*, 31:1539-42

Fréon, P., 1984. La variabilité des tailles individuelles a l'intérieur des cohortes et des bancs de poissons. 1. Observations et interprétation. *Oceanol.Acta*, 7(4):457-68

Fry, F.E.J., 1949. Statistics of a lake trout fishery. *Biometrics*, 5:27-67

Garcia, S., 1985. Reproduction, stock assessment models and population parameters in exploited penaeid shrimp populations. *In* Second Australian national prawn seminar, edited by P.C. Rothlisberg, B.J. Hill and D. Staples. Cleveland, Queensland, Australia, NPS2. :139-58

Garcia, S., 1988. Tropical penaeid prawns. *In* Fish population dynamics: the implications for management, edited by J.A. Gulland, Chichester, John Wiley and Sons Ltd., pp. 219-49

Garcia, S. and L. Le Reste, 1981. Life cycles, dynamics, exploitation and management of coastal penaeid shrimp stocks. *FAO Fish.Tech.Pap.*, (203):215 p. Issued also in French and Spanish.

Garcia, S., P. Sparre and J. Csirke, 1987. A note on rough estimators of fisheries resources potential. *ICLARM Fishbyte*, 5(2):11-6

Garcia, S., P. Sparre and J. Csirke, 1989. Estimating surplus production and maximum sustainable yield from biomass data when catch and effort time series are not available. *Fish.Res.*, 7:

Garcia, S. and N.P. van Zalinge, 1982. Shrimp fishing in Kuwait: methodology for a joint analysis of the artisanal and industrial fisheries. *In* Report on the Workshop on assessment of the shrimp stocks of the west coast of the Gulf between Iran and the Arabian Peninsula. Fisheries development in the Gulf. Rome, FAO, FI:DP/RAB/80/015/1: 119-42

Garrod, D.J., 1961. The selection characteristics of nylon gill nets for *Tilapia esculenta* Graham. *J.Cons.CIEM*, 26(2):191-203

Garrod, D.J., 1988. North Atlantic cod: fisheries and management to 1986. *In* Fish population dynamics: the implications for management, edited by J.A. Gulland, Chichester, John Wiley and Sons Ltd., pp. 185-218

Gavaris, S. and S.J. Smith, 1987. Effect of allocation and stratification strategies on precision of survey abundance estimates for Atlantic cod (*Gadus morhua*) on the eastern Scotian Shelf. *J.Northwest Atl.Fish.Sci.*, 7:137-44

Gayanilo, F.C., Jr., M. Soriano and D. Pauly, 1988. A draft guide to the COMPLEAT ELEFAN. *ICLARM Software Project 2.*: 65 p. and 10 diskettes (5.25 inches, 360K)

Gerking, S.D., 1953. Evidence for concepts of home range and territory in stream fishes. *Ecology*, 34:347-65

Gislason, H. and Th. Helgason, 1985. Species interaction in assessment of fish stocks with special application to the North Sea. *Dana*, 5:1-44

Gislason, H. and P. Sparre, 1987. Some theoretical aspects of the implementation of multispecies virtual population analysis in ICES. ICES C.M. 1987/G:51 (mimeo)

Gjøsaeter, J.P. *et al.*, 1983. Ageing tropical fish by primary growth rings in their otoliths. Bergen, University of Bergen, 67 p. Also issued as:

Gjøsaeter, J.P. *et al.*, 1984. Ageing tropical fish by growth rings in the otoliths. *FAO Fish.Circ.*, (776):54 p.

Goeden, G.B., 1978. A monograph of the coral trout, *Plectropomus leopardus* (Lacépède). *Res.Bull.Fish.Serv.Queensl.*, (1):42 p.

Grosslein, M.D. and A. Laurec, 1982. Bottom trawl surveys design, operation and analysis. CECAF/ECAF Ser., (81/22):25 p. Issued also in French

Gulland, J.A., 1955. On the estimation of population parameters from marked members. Biometrika, 42:269-70

Gulland, J.A., 1961. The estimation of the effect on catches of changes in gear selectivity. J.Cons.CIEM, 26(2):204-14

Gulland, J.A., 1965. Estimation of mortality rates. Annex to Arctic fisheries working group report ICES C.M. Doc. 3 (mimeo)

Gulland, J.A., 1966. Manual of sampling and statistical methods for fisheries biology. Part 1. Sampling methods. FAO Man.Fish.Sci., (3):87 p. Issued also in French and Spanish

Gulland, J.A., 1969. Manual of methods for fish stock assessment. Part 1. Fish population analysis. FAO Man.Fish.Sci., (4):154 p. Issued also in French and Spanish

Gulland, J.A. (comp.), 1971. The fish resources of the ocean. West Byfleet, Surrey, Fishing News (Books), Ltd., for FAO, 255 p. Rev. ed. of FAO Fish.Tech.Pap., (97):425 p. (1970)

Gulland, J.A., 1975. Manual of methods for fisheries resources survey and appraisal. Part 5. Objectives and basic methods. FAO Fish.Tech.Pap., (145):29 p.

Gulland, J.A., 1983. Fish stock assessment: a manual of basic methods. Chichester, U.K., Wiley Interscience, FAO/Wiley series on food and agriculture, Vol.1:223 p.

Gulland, J.A. (ed.), 1988. Fish population dynamics: the implications for management. Chichester, John Wiley and Sons Ltd., 422 p. 2nd ed.

Gulland, J.A., 1988a. The problems of population dynamics and contemporary fishery management. In Fish population dynamics: the implications for management, edited by J.A. Gulland, Chichester, John Wiley and Sons Ltd., pp. 383-406

Gulland, J.A. and D. Harding, 1961. The selection of Clarias mossambicus (Peters) by nylon gill nets. J.Cons.CIEM, 26:215-22

Gulland, J.A. and S.J. Holt, 1959. Estimation of growth parameters for data at unequal time intervals. J.Cons.CIEM, 25(1):47-9

Gulland, J.A. and B.J. Rothschild, (eds), 1984. Penaeid shrimps - their biology and management. Farnham, Surrey, England, Fishing News Books Limited, 308 p.

Gunderson, D.R. and P.H. Dygert, 1988. Reproductive effort as a predictor of natural mortality rate. J.Cons.CIEM, 44:200-9

Hald, A., 1952. Statistical theory with engineering applications. New York, Wiley, 783 p.

Hamley, J.M., 1975. Review of gillnet selectivity. J.Fish.Res.Board Can., 32:1943-69

Hamley, J.M. and H.A. Regier, 1973. Direct estimates of gillnet selectivity to walleye (Stizostedion vitreum vitreum). J.Fish.Res.Board Can., 30:817-30

Hanumara, R.C. and N.A. Hoenig, 1987. An empirical comparison of a fit of linear and non-linear models for seasonal growth in fish. Fish.Res., 5:359-81

Harden Jones, F.R., 1968. Fish migration. London, Arnold, 325 p.

Harden, Jones, F.R., 1984. A view from the ocean. In Mechanisms of migration in fishes, edited by J.D. McCleave et al. New York, Plenum Press, pp. 1-26

Harding, J.P., 1949. The use of probability paper for the graphical analysis of polymodal frequency distributions. J.Mar.Biol.Assoc.U.K., 28:141-53

Hasselblad, V., 1966. Estimation of parameters for a mixture of normal distributions. Technometrics, 8:431-44

Hasselblad, V. and P.K. Tomlinson, 1971. NORMSEP. Normal distribution separator. In Computer programs for fish stock assessment compiled by N.J. Abramson. FAO Fish.Tech.Pap., (101):11(1)2.1-11(1)2.10

Heincke, F., 1913. Investigations on the plaice. General report. 1. The plaice fishery and protective regulations. Part. I. Rapp.P.-V.Réun.-CIEM, 17A:1-153 + Annexes

Helgason, T. and H. Gislason, 1979. VPA-analysis with species interaction due to predation. ICES C.M. 1979/G:52:10 p. (mimeo)

Hoenig, J.M, 1983. Empirical use of longevity data to estimate mortality rates. Fish.Bull.NOAA/NMFS, 81(4):898-903

Holden, M.J. and D.F.S. Raitt (eds), 1974. Manual of fisheries science. Part 2. Methods of resource investigations and their application. FAO Fish.Tech.Pap., (115)Rev.1:214 p. Issued also in French and Spanish

Holland, K. et al., 1985. A small vessel technique for tracking pelagic fish. Mar.Fish.Rev., 47(4):26-32

Holt, S.J., 1962. The application of comparative population studies to fishery biology - an exploration. In The exploitation of natural animal populations, British Ecological Society Symposium Number Two, edited by E.D. Le Cren and M.W. Holdgate. Oxford, Blackwell Scientific Publications pp:51-69

Holt, S.J., 1963. A method for determining gear selectivity and its application. ICNAF Spec.Publ., (5):106-15

Holt, S.J., 1965. A note on the relationship between mortality rate and the duration of life in an exploited fish population. ICNAF Res.-Bull., (2):73-5

Holthuis, L.B., 1980. FAO species catalogue. Vol. 1. Shrimps and prawns of the world. An annotated catalogue of species of interest to fisheries. FAO Fish.Synop., (125)Vol.1:261 p.

Hoydal, K., C.J. Rørvik and P. Sparre, 1980. A method for estimating the effective mesh size and the effect of changes in gear parameters. ICES C.M. 1980/G:28:33 p. (mimeo)

Hoydal, K., C.J. Rørvik and P. Sparre, 1982. Estimation of effective mesh sizes and their utilization in assessment. Dana, 2:69-95

Hunter, J.R. et al., 1986. The dynamics of tuna movements: an evaluation of past and future research. FAO Fish.Tech.Pap., (277):78 p.

I-ATTC (Inter-American Tropical Tuna Commission/Comisión Interamericana del Atún Tropical), 1984. Annual report of the Inter-American Tropical Tuna Commission. Informe anual de la Comisión Interamericano del Atun Tropical, 1983. *Annu.Rep.I-ATTC/Inf.Anu.CIAT*, (1983):272 p.

ICES, 1979. Report of the North Sea round fish working group, Charlottenlund, 7-11 May 1979. ICES, C.M. 1979/G:7 (mimeo)

ICES, 1980. Report of the North Sea round fish working group, Copenhagen, 14-18 April 1980. ICES, C.M. 1980/G:8 (mimeo)

ICES, 1981. Report of the *Ad hoc* working group on the use of effort data in assessment, Copenhagen, 2-6 March 1981. ICES C.M. 1981/G:5 (mimeo)

ICES, 1984. Report of the *Ad hoc* multispecies assessment working group, Copenhagen, 18-22 June 1984. ICES C.M. 1984/Assess:20:99 p. (mimeo)

ICES, 1986. Report of the *Ad hoc* multispecies assessment working group, Copenhagen, 13-19 November 1985. ICES. C.M. 1986/Assess:9:141 p. (mimeo)

ICES, 1987. Report of the *Ad hoc* multispecies assessment working group, Copenhagen, 12-18 November 1986. ICES C.M. 1987/Assess:9 (mimeo)

Ihssen, P.E. *et al.*, 1981. Stock identification: materials and methods. *Can.J.Fish.Aquat.Sci.*, 38:1838-55

Ingles, J. and D. Pauly, 1984. An atlas of the growth, mortality and recruitment of Philippine fishes. *ICLARM Tech.Rep.*, (13):127 p.

Isarankura, A.P., 1971. Assessment of stocks of demersal fish off the west coast of Thailand and Malaysia. FAO/UNDP, Indian Ocean Programme, Rome, IOC/DEV/71/20:20 p.

Jamieson, G.S. and N. Bourne (eds), 1986. North Pacific Workshop on stock assessment and management of invertebrates. *Can.Spec. Publ.Fish. Aquat.Sci*, (92):430 p.

Jensen, J.W., 1986. Gillnet selectivity and the efficiency of alternative combinations of mesh sizes for some freshwater fish. *J.Fish.Biol.*, 28:637-46

Jones, R., 1961. The assessment of the long term effects of changes in gear selectivity and fishing effort. *Mar.Res.Scot.*, 2:19 p.

Jones, R., 1963. Some theoretical observations on the escape of haddock from a codend. *ICNAF Spec.Publ.*, (5):116-127

Jones, R., 1974. Assessing the long term effects of changes in fishing effort and mesh size from length composition data. ICES C.M. 1974/F: 33:13 p. (mimeo)

Jones, R., 1976. Mesh regulation in the demersal fisheries of the South China Sea area. Manila, South China Sea Fisheries Development and Coordinating Programme, SCS/76/WP/34: 75 p.

Jones, R., 1977. Tagging: theoretical methods and practical difficulties. *In* Fish population dynamics, edited by J.A. Gulland. New York, Wiley Interscience, pp. 46-66

Jones, R., 1984. Assessing the effects of changes in exploitation pattern using length composition data (with notes on VPA and cohort analysis). *FAO Fish.Tech.Pap.*, (256):118 p.

Jones, R. and N.P. van Zalinge, 1981. Estimates of mortality rate and population size for shrimp in Kuwait waters. Kuwait Bull.Mar.Sci., 2:273-88

Juanicó, M., 1983. Squid maturity scales for population analysis. FAO Fish.Tech.Pap./FAO Doc.Tech.Pêches/FAO Doc.Téc.Pesca, (231):341-78

Julien-Flüs, M., 1988. A study of growth parameters and mortality rates of *Scomberomorus brasiliensis* from coastal areas of Trinidad, West Indies. FAO Fish.Rep., (389):385-400

Karlsen, L. and B.A. Bjarnasson, 1986. Small-scale fishing with driftnets. FAO Fish.Tech.Pap., (284):64 p.

Kimura, D.K., 1977. Logistic model for estimating selection ogives from catches of codends whose ogives overlap. J.Cons.CIEM, 38(1):116-9

Kleiber, P.M., A.W. Argue and R.E. Kearney, 1983. Investigation of skipjack stock for management purposes. Collect.Vol.Sci.Pap.ICCAT/Recl. Doc.Sci.CICTA/Colecc.Doc.Cient. CICAA 18(3):794-811

Klima, E.F., 1976. An assessment of the fish stocks and fisheries of the Campeche Bank. WECAF Stud., (5):24 p.

Lablache, G. and G. Carrara, 1988. Population dynamics of emperor red snapper, *Lutjanus sebae*, with notes on the demersal fishery on the Mahé Plateau, Seychelles. FAO Fish.Rep., (389):171-92

Lange, A.M.T. and M.P. Sissenwine, 1983. Squid resources of the Northwest Atlantic. FAO Fish.Tech.Pap./FAO Doc.Tech.Pêches/FAO Doc.Téc.Pesca, (231):21-54

Larkin, P.A., 1972. The stock concept and management of Pacific salmon. H.R. MacMillan lectures in fisheries. Vancouver, University of British Columbia Press, 231 p.

Larkin, P.A., 1977. An epitaph for the concept of maximum sustainable yield. Trans.Am.Fish.Soc., 106(1):1-11

Larkin, P.A., 1988. Pacific salmon. In Fish population dynamics: the implications for management, edited by J.A. Gulland, Chichester, John Wiley and Sons Ltd., pp. 153-83

Losse, G.F. and A. Dwippongo, 1977. Report on the Java Sea south east monsoon trawl survey, June-December 1976. Jakarta, Marine Fisheries Research Report (Special report). Contribution of the Demersal Fisheries Project No. 3:119 p.

Ludwig, D., 1987. Estimation of delay-difference model parameters from catch-effort data. Coop.Fish.Res.Unit.Rep.Ser.Univ.B.C.

Ludwig, D. and C.J. Walters, 1985. Are age-structured models appropriate for catch-effort data? Can.J.Fish.Aquat.Sci., 42:1066-72

Macdonald, P.D.M. and T.J. Pitcher, 1979. Age-groups from size-frequency data: a versatile and efficient method of analyzing distribution mixtures. J.Fish.Res.Board Can., 36:987-1001

Mackett, D.J., 1973. Manual of methods for fisheries resource survey and appraisal. Part 3. Standard methods and techniques for demersal fisheries resource surveys. FAO Fish.Tech.Pap., (124):39 p.

MacLean, J.A. and D.O. Evans, 1981. The stock concept, discreteness of fish stocks and fisheries management. Can.J.Fish.Aquat.Sci., 38:1889-98

Mahon, R. (ed.), 1987. Report and proceedings of the expert consultation on shared fishery resources of the Lesser Antilles region. Mayaguez, Puerto Rico, 8-12 September 1986. FAO Fish.Rep., (383):278 p.

Mahon, R. and S. Mahon, 1987. Seasonality and migration of pelagic fishes in the eastern Caribbean. FAO Fish.Rep., (383):192-273

Marten, G.G., 1978. Calculating mortality rates and optimum yields from length samples. J.Fish.Res.Board.Can., 35(2):197-201

Marten, G.G. and J.J. Polovina, 1982. A comparative study of fish yields from various tropical ecosystems. ICLARM Conf.Proc., (9):255-85

Mather, F.J., H.L. Clark and J.M. Mason, 1975. Synopsis of the biology of the white marlin Tetrapterus albidus Poey (1861). NOAA Tech.Rep.NMFS (Spec.Sci.Rep.-Fish.Ser.), (675)Pt.3, Species Synopses:55-94

Mather, F.J., J.M. Mason and H.L. Clark, 1974. Migration of white marlin and blue marlin in the western north Atlantic Ocean - tagging results since May, 1970. NOAA Tech.Rep.NMFS (Spec.Sci.Rep.-Fish.Ser.), (675) Pt.2 Review and contributed papers:211-22

Mather, et al., 1974. Results of sailfish tagging in the western north Atlantic Ocean. NOAA Tech.Rep.NMFS (Spec.Sci.Rep.-Fish.Ser.), (675)Pt.2 Review and contributed papers:194-210

Matsumoto, W.M., R.A. Skillman and A.E. Dizon, 1984. Synopsis of biological data on skipjack tuna, (Katsuwonus pelamis). FAO Fish.Synop., (136):92 p. Issued also as NOAA Tech.Rep.NMFS Circ., (451):92 p.

McCombie, A.M. and F.E.J. Fry, 1960. Selectivity of gill nets for lake whitefish Coregonus clupeaformis. Trans.Am.Fish.Soc., 89:176-84

McNew, R.W. and R.C. Summerfelt, 1978. Evaluation of a maximum-likelihood estimator for analysis of length-frequency distributions. Trans.Am.Fish.Soc., 107(5):730-6

Mesnil, B., 1988. Logiciels pour l'évaluation des stocks de poisson. ANACO: logiciel d'analyse des données de capture par classes d'âge sur IBM PC et compatibles. FAO Doc.Tech.Pêches, (101)Suppl.3:78 p.

Mohamed, K.H., et al., 1979. Mark-recapture experiments on the Gulf shrimp, Penaeus semisulcatus, De Haan, in Kuwait waters. Shrimp stock evaluation and management project, Kuwait. Report TF/KUW-6/R-10:59 p. (mimeo)

Morales-Nin, B.M. (in press). Determinación del crecimiento de peces teleósteos tropicales por media de la microestructura de los otolitos y de los análisis de frecuencia de tallas. [Determination of growth of tropical bony fish by means of micro-structures in the otoliths and length-frequency analyses]. FAO Doc.Téc.Pesca (in prensa). (Will also be translated into English)

Moreau, J., C. Bambino and D. Pauly, 1986. Indices of overall fish growth performance of 100 tilapia (Cichlidae) populations. In The first Asian fisheries forum, edited by J.L. Maclean, L.B. Dizon and L.V. Hosillos, Manila, Philippines, Asian Fisheries Society: pp. 201-6

Munro, J.L., 1974. The mode of operation of Antillean fish traps and the relationships between ingress, escapement, catch and soak. J.Cons. CIEM, 35(3):337-50

Munro, J.L., (ed.), 1983. Caribbean coral reef fishery resources. ICLARM Stud.Rev., (7):276 p.

Munro, J.L. and D. Pauly, 1983. A simple method for comparing growth of fishes and invertebrates. ICLARM Fishbyte, 1(1):5-6

Munro, J.L. and R. Thompson, 1983. The Jamaican fishing industry. ICLARM Stud.Rev., (7):10-4

Munro, J.L. and R. Thompson, 1983a: Areas investigated, objectives and methodology. ICLARM Stud.Rev., (7):15-25

Murawski, S.A., 1984. Mixed species yield-per-recruitment analyses accounting for technological interactions. Can.J.Fish.Aquat.Sci., 41:897-916

Murphy, G.I., 1966. Population biology of the pacific sardine (Sardinops caerulea). Proc.Calif.Acad.Sci., 34:1-84

Murphy, G.I., 1977. Clupeoids. In Fish population dynamics, edited by J.A. Gulland. New York, Wiley Interscience, pp. 283-308

Murphy, G.I., 1982. Recruitment of tropical fishes. ICLARM Conf.Proc., (9):141-8

Naamin, N. and S. Noer, 1980. The status of the shrimp fishery in the Arafura Sea. In Report of the Workshop on the biology and resources of penaeid shrimps in the South China Sea area. Part 1. Manila, South China Sea Fisheries Development and Coordinating Programme, SCS/GEN/-80/26:13-32

Nédélec, C., 1982. Definition and classification of fishing gear categories. FAO Fish.Tech.Pap., (222):51 p. Issued also in French and Spanish

Nishikawa, Y., et al., 1985. Average distribution of larvae of oceanic species of Scrombroid fishes, 1956-1981. Far Seas Fish.Res.Lab., SSeries(12):99 p.

Osako, M. and M. Murata, 1983. Stock assessment of cephalopod resources in the Northwestern Pacific. FAO Fish.Tech.Pap./FAO Doc.Tech.Pêches/FAO Doc.Téc.Pesca, (231):55-144

Oxenford, H.A. and W. Hunte, 1986. Migration of the dolphin (Coryphaena hippurus) and its implications for fisheries management in the Western Central Atlantic. Proc.Gulf.Caribb.Fish.Inst., 37:95-111

Packard, A., 1966. Optional convergence between cephalopods and fish: an exercise in functional anatomy. Arch.Zool.Ital., 51:523-43

Paloheimo, J.E., 1958. A method of estimating natural and fishing mortalities. J.Fish.Res.Board Can., 15(4):749-58

Paloheimo, J.E., 1961. Studies on estimation of mortalities. I. Comparison of a method described by Beverton and Holt and a new linear formula. J.Fish.Res.Board Can., 18(5):645-62

Paloheimo, J.E., 1980. Estimation of mortality rates in fish populations. Trans.Am.Fish.Soc., 109:378-86

Paloheimo, J.E. and E. Cadima, 1964. On statistics of mesh selection. ICNAF Serial No. 1394/Doc.No. 98

Panella, G., 1971. Fish otoliths: daily growth layers and periodical patterns. Science, Wash., 137(4002):1124-7

Parrish, B.B., (ed.), 1973. Fish stocks and recruitment. Proceedings of a symposium held in Aarhus 7-10 July 1970. Rapp.P.-V.Réun.CIEM, (164):-372 p.

Paulik, G.J., 1963. Estimates of mortality rates from tag recoveries. Biometrics, 19:28-57

Pauly, D., 1979. Theory and management of tropical multispecies stocks: a review with emphasis on the Southeast Asian demersal fisheries. ICLARM Stud.Rev., (1):35 p.

Pauly, D., 1980: A selection of simple methods for the assessment of tropical fish stocks. FAO Fish.Circ., (729):54 p. Issued also in French. Superseded by FAO Fish.Tech.Pap., (234)

Pauly, D., 1980a: A new methodology for rapidly acquiring basic information on tropical fish stocks: growth, mortality and stock-recruitment relationships. In Stock assessment for tropical small-scale fisheries. Proceedings of an International Workshop held 19-21 September 1979, edited by S.B. Saila and P. Roedel. Kingston, University of Rhode Island, International Center for Marine Resource Development, pp. 154-72

Pauly, D., 1980b: On the interrelationships between natural mortality, growth parameters, and mean environmental temperature in 175 fish stocks. J.Cons.CIEM, 39(2):175-92

Pauly, D., 1981: Tropical stock assessment package for programmable calculators and micro-computers. ICLARM Newsl., 4(3):10-3

Pauly, D., 1983. Some simple methods for the assessment of tropical fish stocks. FAO Fish.Tech.Pap., (234):52 p. Issued also in French and Spanish

Pauly, D., 1983a. Length-converted catch curves. A powerful tool for fisheries research in the tropics. (Part I). ICLARM Fishbyte, 1(2):9-13

Pauly, D., 1984. Fish population dynamics in tropical waters: a manual for use with programmable calculators. ICLARM Stud.Rev., (8):325 p.

Pauly, D., 1984a. Length-converted catch curves. A powerful tool for fisheries research in the tropics. (Part II). ICLARM Fishbyte, 2(1):17-9

Pauly, D., 1984b. Length-converted catch curves. A powerful tool for fisheries research in the tropics. (III: Conclusion). ICLARM Fishbyte, 2(3):9-10

Pauly, D., 1987. A review of the ELEFAN system for analysis of length-frequency data in fish and aquatic invertebrates. ICLARM Conf.Proc., (13):7-34

Pauly, D., 1988. Fisheries research and the demersal fisheries of Southeast Asia. In Fish population dynamics: the implications for management, edited by J.A. Gulland, Chichester, John Wiley and Sons Ltd., pp. 329-48

Pauly, D. and J.F. Caddy, 1985. A modification of Bhattacharya's method for the analysis of mixtures of normal distributions. FAO Fish.Circ., (781):16 p.

Pauly, D. and N. David, 1981. ELEFAN I, a BASIC program for the objective extraction of growth parameters from length-frequency data. Meeresforschung, 28(4):205-11

Pauly, D. and G. Gaschütz, 1979. A simple method for fitting oscillating length growth data, with a program for pocket calculators. ICES C.M. 1979/G:24:26 p. (mimeo)

Pauly, D. and G.R. Morgan (eds), 1987. Length-based methods in fisheries research. ICLARM Conf.Proc., (13):468 p.

Pauly, D. and J.L. Munro, 1984. Once more on the comparison of growth in fish and invertebrates. ICLARM Fishbyte, 2(1):21

Pauly, D. and G.I. Murphy (eds) 1982. Theory and management of tropical fisheries. ICLARM Conf.Proc., (9):360 p.

Pauly, D. and N.A. Navaluna, 1983. Monsoon-induced seasonality in the recruitment of Philippine fishes. FAO Fish.Rep., (291)Vol.3:823-33

Pearson, K., 1894. Contribution to the mathematical theory of evolution. Philos.Trans.R.Soc.Lond.(Ser.A.), 185:71-110

Penn, J.W., 1984. The behavior and catchability of some commercially exploited penaeids and their relationship to stock and recruitment. In Penaeid shrimps - their biology and management, edited by J.A. Gulland and B.J. Rothschild. Farnham, Surrey, England, Fishing News Books Ltd., pp. 173-86

Pennington, M., 1983. Efficient estimators of abundance, for fish and plankton surveys. Biometrics, 39:281-6

Petersen, C.G.J., 1892. Fiskenes biologiske forhold i Holbaek Fjord, 1890-91. Beret.Danm.Biol.St., 1890(1)1:121-183. (in Danish)

Peterson, I. and S.J. Wroblewski, 1984. Mortality rate of fishes in the pelagic ecosystem. Can.J.Fish.Aquat.Sci., 41:1117-20

Pitcher, T.J. and P.D.M. Macdonald, 1973. Two models for seasonal growth in fishes. J.Appl.Ecol., 10:597-606

Polovina, J.J., 1984. An overview of the ECOPATH model. ICLARM Fishbyte, 2(2):5-7

Polovina, J.J., 1984a. Model of the coral reef ecosystem. Part I. The ECOPATH model and its application to French frigate shoals. Coral Reefs, 3(1):1-11

Pope, J.A., et al., 1975. Manual of methods for fish stock assessment. Pt3. Selectivity of fishing gear. FAO Fish.Tech.Pap., (41)Rev.1:65 p.

Pope, J.G., 1972. An investigation of the accuracy of virtual population analysis using cohort analysis. Res.Bull.ICNAF, (9):65-74

Pope, J.G., 1979. Stock assessment in multispecies fisheries with special reference to the trawl fisheries in the Gulf of Thailand. Manila. South China Sea Development and Coordinating Programme, SCS/DEV/79/-19:106 p.

Pope, J.G., 1979a. A modified cohort analysis in which constant natural mortality is replaced by estimates of predation levels. ICES C.M. 1979/H:16:7 p. (mimeo)

Pope, J.G., 1980. Assessment of multispecies resources. In Selected lectures from the CIDA/FAO/CECAF seminar on fishery resource evaluation. Casablanca, Morocco, 6-24 March 1978. Rome, FAO, Canada Funds-in-Trust, FAO/TF/INT 180(c) Suppl.:93-137

Pope, J.G., 1980a. Phalanx analysis: an extension on Jones' length cohort analysis to multispecies cohort analysis. ICES C.M. 1980/G:19:18 p. (mimeo)

Pope, J.G., 1988. Collecting fisheries assessment data. In Fish population dynamics: the implications for management, edited by J.A. Gulland, Chichester, John Wiley and Sons Ltd., pp. 63-82

Postel, E., 1955. Contribution a l'étude de la biologie de quelques Scombridae de l'Atlantique tropico-oriental. Ann.Stn.Oceanogr.Salammbo, (10):167 p.

Pütter, A., 1920. Studien über physiologische Ähnlichkeit. VI. Wachstumsähnlichkeiten. Pflüger Arch.Ges.Physiol., 180:298-340

Raj, D., 1968. Sampling theory. New York, McGraw-Hill.

Ralston, S., 1982. Influence of hook size in the Hawaiian deep-sea handline fishery. Can.J.Fish.Aquat.Sci., 39:1297-302

Randall, J.E., 1962. Tagging reef fishes in the Virgin Islands. Proc.Gulf Caribb.Fish.Inst., 14:201-41

Rao, C.R., 1965. Linear statistical interference and its applications. New York, John Wiley, 625 p. 2nd ed.

Regier, H.A. and D.S. Robson, 1966. Selectivity of gill nets, especially to lake whitefish. J.Fish.Res.Board Can., 23:423-54

Ricker, W.E., 1954. Stock and recruitment. J.Fish.Res.Board Can., 11:559-623

Ricker, W.E., 1973. Linear regressions in fishery research. J.Fish.Res.-Board.Can., 30:409-34

Ricker, W.E., 1975. Computation and interpretation of biological statistics of fish populations. Bull.Fish.Res.Board Can., (191):382 p.

Riedel, D., 1963. Contribution to the experimental determination of the selection parameter of gill nets. Arch.Fischereiwiss., 14:85-97

Rikhter, V.A. and V.N. Efanov, 1976. On one of the approaches to estimation of natural mortality of fish populations. ICNAF Res.Doc., 76/VI/8:12 p.

Robson, D.S., 1966. Estimation of relative fishing power of individual ships. Res.Bull.ICNAF, (3):5-14

Robson, D.S. and D.G. Chapman, 1961. Catch curves and mortality rates. Trans.Am.Fish.Soc., 90(2):181-9

Rodríguez, A., 1977. Contribución al conocimiento de la biologia y pesca del langostino, Penaeus kerathurus (Forskål 1775) del golfo de Cádiz (Región Sudatlántica española). Invest.Pesq.Barc., 41(3):603-35

Roper, C.F.E., M.J. Sweeney and C.E. Nauen, 1984. FAO species catalogue. Vol. 3. Cephalopods of the world. An annotated and illustrated catalogue of species of interest to fisheries. FAO Fish.Synop., (125)Vol.3:277 p.

Rosenberg, A.A. and J.R. Beddington, 1988. Length-based methods of fish stock assessment. In Fish population dynamics: the implications for management, edited by J.A. Gulland, Chichester, John Wiley and Sons Ltd., pp. 83-103

Rothlisberg, P.C., B.J. Hill and D.J. Staples, (eds), 1985. Second Australian national prawn seminar. Cleveland, Queensland. Australia, NPS2:361 p.

Rudstam, L.G., J.J. Magnuson and W.M. Tonn, 1984. Size selectivity of passive fishing gear: a correction for encounter probability applied to gill nets. Can.J.Fish.Aquat.Sci., 41:1252-5

Saeger, J., P. Martosubroto and D. Pauly, 1976. First report of the Indonesian-German demersal fisheries project (Result of a trawl survey in the Sunda Shelf area). Jakarta, Marine Fisheries Research Report (Special report). Contribution of the Demersal Fisheries Project no. 1:46 p.

Saila, S.B. and P. Roedel (eds), 1980. Stock assessment for tropical small-scale fisheries. Proceedings of an International Workshop held 19-21 September 1979, at the University of Rhode Island, Kingston R.I., International Center for Marine Resources Development, University of Rhode Island, 198 p.

Sainsbury, K.J., 1988. The ecological basis of multispecies fisheries, and management of a demersal fishery in tropical Australia. In Fish population dynamics: the implications for management, edited by J.A. Gulland, Chichester, John Wiley and Sons Ltd., pp. 349-82

Sato, T. and H. Hatanaka, 1983. A review of assessment of the Japanese distant-water fisheries for cephalopods. FAO Fish.Tech.Pap./FAO Doc.Tech.Pêches/FAO Doc.Téc.Pesca, (231):145-80

Saville, A. (ed.), 1977. Survey methods of appraising fisheries resources. FAO Fish.Tech.Pap., (171):76 p.

Saville, A. (ed.), 1980. The assessment and management of pelagic fish stocks. A symposium held in Aberdeen, 3-7 July 1978. Rapp.P.-V.Réun. CIEM, 177:517 p.

Schaefer, M., 1954. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. Bull.I-ATTC/Bol.CIAT, 1(2):27-56

Schaefer, M., 1957. A study of the dynamics of the fishery for yellowfin tuna in the eastern tropical Pacific Ocean. Bull.I-ATTC/Bol.CIAT, 2:247-68

Schnute, J., 1985. A general theory for analysis of catch and effort data. Can.J.Fish.Aquat.Sci., 42:414-29

Schnute, J., 1987. A general fishery model for a size-structured fish population. Can.J.Fish.Aquat.Sci., 44(5):924-40

Schnute, J. and D. Fournier, 1980. A new approach to length-frequency analysis: growth structure. Can.J.Fish.Aquat.Sci., 37:1337-51

SCSP (South China Sea Development Programme), 1978. Report on the workshop on the demersal resources of the Sunda Shelf, Part 1. Manila, South China Sea Fisheries Development and Coordinating Programme, SCS/GEN/77/12:44 p.

Seber, G.A.F., 1973. The estimation of animal abundance and related parameters. London, Griffin, 506 p.

Sharp, G.D., 1978. Behavioral and physiological properties of tunas and their effects on vulnerability to fishing gear. In The physiological ecology of tunas, edited by G.D. Sharp and A.E. Dizon. New York, Academic Press, pp. 397-449

Sharp, G.D. and A.E. Dizon (eds), 1978. The physiological ecology of tunas. New York, Academic Press, 485 p.

Shepherd, J.G., 1988. Fish stock assessments and their data requirements. In Fish population dynamics: the implications for management, edited by J.A. Gulland, Chichester, John Wiley and Sons Ltd., pp. 35-62

Shindo, S., 1973. General review of the trawl fishery and the demersal fish stocks of the South China Sea. FAO Fish.Tech.Pap., (120):49 p.

Sickle, van J., 1977. Mortality rates from size distributions: the application of a conservation law. Oecologia, Berl., 27:311-8

Sissenwine, M.P., 1978. Is MSY an adequate foundation for optimum yield? Fisheries, 3(6):22-4 and 37-42

Sissenwine, M.P., M.J. Fogarty and W.J. Overholtz, 1988. Some fisheries management implications of recruitment variability. In Fish population dynamics: the implications for management, edited by J.A. Gulland, Chichester, John Wiley and Sons Ltd., pp. 129-52

Smith, T.D., 1988. Stock assessment methods: the first fifty years. In Fish population dynamics: the implications for management, edited by J.A. Gulland, Chichester, John Wiley and Sons Ltd., pp. 1-33

Sokal, R.R. and F.J. Rohlf, 1981. Biometry. The principles and practice of statistics in biological research. San Francisco, California, Freeman and Company, 2nd ed.

Som, R.K., 1973. A manual of sampling techniques. London, Heinemann Educational Books Ltd., 384 p.

Sousa, M.I., 1988. Sources of bias in growth and mortality estimation of migratory pelagic fish stocks, with emphasis on Decapterus russelli (Carangidae) in Mozambique. FAO Fish.Rep., (389):288-307

Spain, J.D., 1982. BASIC microcomputer model in biology. Reading, Massachusetts. Addison-Wesley Publishing Company, 354 p.

Sparre, P., 1979. Some necessary adjustments for using the common methods in eel assessment. Rapp.P.-V.Réun.CIEM, 174:41-44

Sparre, P., 1979. Some remarks on the application of yield/recruit curves in estimation of maximum sustainable yield. ICES C.M. 1979/G:41:21 p. (mimeo)

Sparre, P., 1980. A goal function of fisheries (Legion analysis). ICES C.M. 1980/G:40:81 p. (mimeo)

Sparre, P., 1983. Legion analysis game and simulation program (LAGS). NAFO SCR Doc. 83/IX/71/Ser.No. N736:49 p.

Sparre, P., 1984. A computer program for estimation of food suitability coefficients from stomach content data and multispecies VPA. ICES C.M. 1984/G:25:60 p. (mimeo)

Sparre, P., 1987. Computer programs for fish stock assessment. Length-based fish stock assessment for Apple II computers. FAO Fish.Tech. Pap., (101) Suppl.2:218 p.

Sparre, P., 1987a. A method for the estimation of growth, mortality and gear selection/recruitment parameters from length-frequency samples weighted by catch per effort ICLARM Conf.Proc., (13):75-102

SPC (South Pacific Commission), 1984. An assessment of the skipjack and baitfish resources of northern Mariana Islands, Guam, Palau, Federated States of Micronesia and Marchal Islands. Final Ctry Rep.Skipjack Surv.Assess. Programme S.Pac.Comm.,(18):111 p.

Stéquert, B. and F. Marsac, 1986. La pêche de surface de thonidés tropicaux dans l'océan Indien. FAO Doc.Tech.Pêches, (282):213 p.

Sturm, M.G. de L., 1974. Aspects of the biology of the Spanish mackerel, Scomberomorus maculatus (Mitchell) in Trinidad, West Indies. Trinidad, University of the West Indies, Ph.D. Thesis

Sukhatme, R.V. and B.V. Sukhatme, 1970. Sampling theory of surveys with applications. Ames, Iowa, Iowa State University Press, 452 p.

Sund, P.N., M. Blackburn and F. Williams, 1981. Tunas and their environment in the Pacific Ocean: A review. Oceanogr.Mar.Biol., 19:443-512

Tanaka, S., 1953. Precision of age-composition of fish estimated by double sampling method using the length for stratification. Bull.Jap.Soc. Sci.Fish., 19:657-70

Tanaka, S., 1960. Studies on the dynamics and the management of fish populations. Bull.Tokai.Reg.Fish.Res.Lab., (28):1-200 (In Japanese)

Taylor, C.C., 1960. Temperature, growth and mortality - the Pacific cockle. J.Cons.CIEM, 26:117-24

Tiews, K. and P. Caces-Borja, 1965. On the availability of fish of the family Leiognathidae Lacépedè in Manila Bay and San Miguel Bay and on their accessibility to controversial fishing gears. Philipp.J.Fish., 7(1):59-85

Tiurin, P.V., 1962. The natural mortality factor and its importance in regulating fisheries. Vopr.Ikhtiol., (2):403-27 (in Russian)

Thompson, D.W., 1910. The works of Aristotele. Vol. 4. Oxford, Clarendon Press.

Thompson, W.F. and F.H. Bell, 1934. Biological statistics of the Pacific halibut fishery. 2. Effect of changes in intensity upon total yield and yield per unit of gear. Rep.Int.Fish. (Pacific Halibut) Comm., (8):49 p.

Troadec, J.-P., 1977. Méthodes semi-quantitatives d'évaluation. FAO Circ. Pêches, (701):131-141

Troadec, J.-P., 1980. Utilization of resource survey results in stock assessment. In Selected lectures from the CIDA/FAO/CECAF seminar on fishery resource evaluation. Casablanca, Morocco, 6-24 March 1978. Rome, FAO, Canada Funds-in-Trust, FAO/TF/INT 180(c) Suppl.:139-52

Ursin, E., 1968. A mathematical model of some aspects of fish growth, respiration and mortality. J.Fish.Res.Board Can., 24:2355-453

Ursin, E., 1982. Stability and variability in the marine ecosytem. Dana, 2:51-67

Ursin, E., 1984. The tropical, the temperate and the arctic seas as media for fish production. Dana, 3:43-60

Vakily, J.M., M.L. Palomares and D. Pauly, 1986. Computer programs for fish stock assessment: Applications for the HP 41 CV calculator. Produced with the cooperation of the International Center for Living Aquatic Resources Management (ICLARM). FAO Fish.Tech.Pap., (101)Suppl.1:255 p.

Venema, S.C., J.M. Christensen and D. Pauly, (eds), 1988. Contributions to tropical fisheries biology. Papers prepared by the participants at the FAO/DANIDA Follow-up Training Courses on fish stock assessment in the tropics. Hirtshals, Denmark, 5-30 May 1986 and Manila, Philippines, 12 January-6 February 1987. FAO Fish.Rep., (389):519 p.

Venema, S.C., J.M. Christensen and D. Pauly, 1988. Training in tropical fish stock assessment: a narrative of experience. FAO Fish.Tech.Pap., (389):1-15

Walford, L.A., 1946. A new graphic method of describing the growth of animals. Biol.Bull.Mar.Biol.Lab.Woods Hole, 90:141-7

Weber, W. and A.A. Jothy, 1977. Observations on the fish Nemipterus spp. (Family: Nemipteridae) in the coastal waters of East Malaysia. Arch.Fischereiwiss., 28(2/3):109-22

Wetherall, J.A., J.J. Polovina and S. Ralston, 1987. Estimating growth and mortality in steady-state fish stocks from length-frequency data. ICLARM Conf.Proc., (13):53-74

Whitehead, P.J.P., 1985. FAO species catalogue. Vol. 7. Clupeoid fishes of the world. An annotated and illustrated catalogue of the herrings, sardines, pilchards, sprats, anchovies and wolfherrings. Part 1. Chirocentridae, Clupeidae and Pristigasteridae. FAO Fish.Synop., (125)Vol.7, Pt.1:303 p.

Whitehead, P.J.P., G.J. Nelson and T. Wongratana, 1988. FAO species catalogue. Vol. 7. Clupeoid fishes of the world (suborder Clupeoidei). An annotated and illustrated catalogue of the herrings, sardines, pilchards, sprats, anchovies and wolfherrings. Part 2. Engraulididae. FAO Fish. Synop., (125)Vol.7.Pt.2.:305-579 p.

Williams, T.P., 1986. Ageing manual for Kuwait fish. KISR 1915. MB-44. Kuwait Institute for Scientific Research.

Willmann, R. and S.M. Garcia, 1985. A bio-economic model for the analysis of sequential artisanal and industrial fisheries for tropical shrimp (with a case study of Suriname shrimp fisheries). FAO Fish.Tech.Pap., (270):49 p. Issued also in French and Spanish

Worms, J., 1983. World fisheries for cephalopods: a synoptic overview. FAO Fish.Tech.Pap./FAO Doc.Tech.Pêches/FAO Doc.Téc.Pesca, (231):1-20

Wyatt, J.R., 1983. The biology, ecology and bionomics of the squirrelfishes, Holocentridae. ICLARM Stud.Rev., (7):50-8

Yates, G.T., 1983. Hydromechanics of body and caudal fin propulsion. In Fish biomechanics, edited by P.W. Webb and D. Weihs. New York, Praeger Publishers, pp. 177-213

Yonimori, T., 1982. Study of tuna behaviour, particularly their swimming depths by use of sonic tags. Newsl.Enyo (Far Seas) Fish.Res.Lab.Shimizu, (44):1-5 Transl. from Japanese

Yuen, H.S.H., 1970. Behavior of skipjack tuna, *Katsowonus pelamis*, as determined by tracking with ultrasonic devices. J.Fish.Res.Board Can., 27(11):2071-9

Zalinge, N. van and P. Sparre, 1986. Pakistan. Statistical systems and proposed sampling scheme. Rome, FAO, FI:DP/PAK/77/003, Field Document 6

Ziegler, B., 1979. Growth and mortality rates of some fishes of Manila Bay, Philippines, as estimated from analysis of length-frequencies. Thesis. Kiel University, 115 p.

16 SUBJECT INDEX

Subject: . . . . Section(s)

Active gear 6.2
Age 1.6
Age at first capture, Tc. 4.1, 8.1
Age of massive maturation, Tm50 . . . . 4.7.3
Age structured model. 1.3.1
Age/length key. 3.2.1, 3.3.5
Amplitude (ELEFAN), C 3.5.2
Analytical model. 1.3.1
Arbitrary age 3.1.4, 3.4.1
Asymptotic body length, L_∞. 3.1
Asymptotic body weight, W_∞. 3.3.6
Attrition rate, A 11.5
Average number of survivors (from cohort), N. 4.2

Baranov's equation. 4.2
Bertalanffy, von, growth model. 3.1
Bertalanffy, von, growth parameters 3.1
Bertalanffy, von, inverse growth equation 3.3.3
Bertalanffy, von, plot. 3.3.3
Bertalanffy, von, seasonalized growth equation. 3.5.2
Bertalanffy, von, weight converted growth equation. 3.3.6
Beverton and Holt's stock/recruitment model 12.1
Beverton and Holt's Z-equations 4.5
Bhattacharya plot 2.6
Bhattacharya method 3.4.1
Bias. 1.7, 2.2, 2.7, 3.4.1.3
Biological interaction. 10.2
Biomass estimation. 5, 8.3 13.6
Biomass per recruit (Beverton and Holt, B/R 8.3
Birth, day of 1.6, 3.1
Birth, estimation of approximative. 3.4.2
Body depth. 2.4
Bottom trawl. 6.0, 13.1
Bottom trawl survey 13.0

Cadima's formula. 9.4
Cassie's method 3.4.3
Catch curve 4.4.2
Catch curve, cumulated. 4.4.4, 4.4.6
Catch curve, based on age composition (variable time intervals) . . 4.4.4
Catch curve, based on age composition (constant time interval). . . 4.4.3
Catch curve, based on length composition data 4.4.5
Catch equation (see Baranov's equation) 4.2
Catch per unit of area, CPUA. 13.6
Catch per unit of effort, CPUE. 4.3, 9.0
Catchability coefficient, q 4.3, 9.2
Codend. 13.1
Coefficient of variation. 2.1
Cohort. 1.3.1, 4.1
Cohort analysis, age based (Pope's) 5.2
Cohort analysis, length based (Jones'). 5.4
Condition factor, q 3.3.6
Confidence interval 2.3
Confidence limits 2.3
Confidence limits of slope and intercept. 2.4
Confidence limits of correlation coefficient. 2.5
Constant parameter system 4.4.1
Correction for mesh selection 6.5
Correlation coefficient 2.5

CPUA, catch per unit of area 13.6
CPUE, catch per unit of effort 4.3, 9.0
Cumulated catch curve 4.4.4
Cumulated catch curve based on length composition data 4.4.6
Curvature parameter, K 3.1

Data massage 3.4.2
Data recording 13.3
Deck sampling 13.4
Density dependent mortality 12.1
Delta function 13.7
Dependent variable 2.4
Depth (body depth) 2.4
Deriso's and Schnute's delay/difference model 9.9
Distribution, stock 1.1
Dynamics of a cohort 4.2

Economic interaction 10.3
Effective mesh size 6.4
Effort (standardization of effort) 9.8
ELEFAN I 3.5.1
Equilibrium situation 9.2
Exact Gulland and Holt plot 3.3.1
Exploitation rate, E = F/Z 4.2
Exploited phase (of the cohort's life) 4.1
Exploratory survey 13.0
Exponential decay model 4.2, 8.1

F, fishing mortality coefficient 4.2
Finite population correction factor 2.7
Fishing mortality coefficient, F 4.2
Fishing pattern 5.1, 6.6
Fishing power 9.8
Fleet 7.0
Ford-Walford plot 3.3.2
Fork length 1.5
Fox model 9.1
Frequency 2.1
Frequency, theoretical 2.2
Frequency, relative theoretical 2.2
Frequency, relative observed 2.2
Frequency, cumulated relative theoretical 2.2
Frequency, cumulated relative observed 2.2
Frequency table 2.1
Functional regression analysis 2.5

Garcia *et al*. models 9.3
Gear selection curve 4.5, 6.0
Gill net selection 6.2
Gonadal maturity 3.6
Growth overfishing 8.0, 8.2
Growth parameters 1.1, 3.0
Growth rate 3.1
Gulland's formula 9.3
Gulland and Holt plot 3.3.1

Hanging ratio 6.2
Heincke's method 4.3.1
Holistic model 1.3.2
Homing 11.1

Independent variable 2.4
Initial condition parameter (t_o, t-zero) 3.1
Instantaneous rate of fishing mortality, F 4.2

Instantaneous rate of natural mortality, M 4.7
Instantaneous rate of total mortality, Z 4.2
Intercept 2.4
Interval midpoint 2.2
Inverse regression analysis 2.5
Inverse von Bertalanffy growth equation 3.3.3

Jones' length based cohort analysis 5.4
Jones and van Zalinge's method 4.4.6

K (curvature parameter) 3.1
Knife-edge selection 6.1.1, 8.1
Knife-edge recruitment 8.1

Length, body 1.5
Length based Thompson and Bell analysis 5.5
Length based cohort analysis (Jones') 5.4
Length/weight relationship 2.6, 3.3.6
LFSA package 14.0
Linear regression 2.4
L-infinity (L_∞) 3.1
Linearized catch equation 4.4.2
Linearization of a normal distribution 2.6
Logistic curve 6.1.1
Log sheets 13.3
Longevity 4.7.1

M, natural mortality coefficient 4.7
Matched samples method 11.3, 11.4
Maximum fishing mortality 6.1.1
Maximum likelihood methods 3.5.3
Maximum sustainable yield (MSY) 1.1, 8.0, 9.1, 13.8
Maximum sustainable yield, Beverton and Holt 4.5, 8.0
Maximum sustainable yield, Fox 9.1
Maximum sustainable yield, per recruit 8.2
Maximum sustainable yield, Schaefer 9.1
Maximum sustainable yield, Garcia et al. 9.3
Maximum sustainable yield, Deriso/Schnute 9.7
Mean value 2.1
Mean value from frequency sample 2.1
Mesh selection experiments 6.1.3
Mesh size 6.1.3
Migration 1.7, 11.1
Migration coefficient 11.4
Mixed fishery, assessment of 10.4
Modal progression analysis 3.4.2
Model 1.3
Monte Carlo simulation technique 3.5.4
Mortality rates 1.1, 4.2, 4.7
MSY = Maximum Sustainable Yield 1.1, 8.0, 9.1, 13.8
Multifleet mixed fisheries 10.4.3
Multispecies VPA 10.2
Munro and Thompson's plot 9.7

N, number of survivors (from cohort) 4.1
$\bar{N}$, average number of survivors (from a cohort) 4.2
Natural mortality coefficent, M 4.7
Neyman allocation 2.8
Normal distribution 2.2
Normal distribution graph paper 2.2
Normsep 3.5.3
Number of survivors (from a cohort), N 4.1
Number of recruits, R, N(Tr) 4.1

Optimum stratified random sampling allocation . . . 2.8
Overfishing, growth . . . 1.3.1, 8.0
Overfishing, recruitment . . . 1.3.1, 8.3

Passive gear . . . 6.2
Pauly's empirical M-formula . . . 4.7.2
Parabola method . . . 3.4.3
Parameter . . . 1.2, 2.4
Phi prime (ϕ') . . . 3.7
Plot of Z on effort . . . 4.6
Population parameters . . . 1.1
Predation mortality . . . 10.2
Pre-recruitment phase . . . 4.1
Predation mortality . . . 10.2
Probability . . . 2.2
Probability paper . . . 2.2
Probability paper method (Cassie method) . . . 3.4.3
Proportional sampling . . . 2.8
Pseudo-cohort . . . 4.4.1, 5.4

q, catchability coefficient . . . 4.3
q, condition factor . . . 3.3.6

R, number of recruits, N(Tr) . . . 4.1
Raising factors . . . 7.3
Random sampling . . . 1.7, 2.7
Recruit . . . 1.6
Recruitment . . . 1.6, 4.1
Recruitment, age of, Tr . . . 4.1
Recruitment, curve . . . 6.1
Recruitment, intensity . . . 1.6
Recruitment, overfishing . . . 8.1, 8.3
Recruitment, season . . . 1.6
Regression . . . 2.4
Regression line . . . 2.4
Relative yield per recruit (Beverton and Holt) . . . 8.4
Resultant selection curve . . . 6.1
Reversed logistic curve . . . 6.2
Ricker's stock/recruitment model (Ricker curve) . . . 12.1
Rikhter and Efanov's formula . . . 4.7.3
Robson and Chapman's method . . . 4.3.2

Sampling . . . 2.7, 2.8, 7.0, 13.0
Sampling of commercial catches . . . 7.0
Scatter diagram . . . 2.4
Schaefer model . . . 9.1
Schaefer model, multispecies . . . 10.1
Selection curve . . . 6.1
Selection factor . . . 6.1.3
Selection range . . . 6.1.1
Separation index, I . . . 3.5.4
Simple random sampling . . . 2.7
Slope . . . 2.4
Standard deviation . . . 2.1
Standard error . . . 2.3
Standard length . . . 1.5
Standardization of effort . . . 9.8
Stock . . . 1.1
Stock parameters . . . 1.1
Stock/recruitment relationship . . . 12.0
Stratified random sampling . . . 2.8, 13.7
Stratified random sampling, optimum . . . 2.8
Stratum . . . 2.8
Student's t-distribution . . . 2.3

Summerpoint (ELEFAN), ts. 3.5.2
Surplus production models 1.3.2, 9.0
Surplus production models, multispecies 10.1, 10.2
Survival rate 4.2
Swept area. 13.5
Swept area method 1.3.2, 4.3, 13.6

Technical interaction 10.4
Thompson and Bell model, age based. 5.3
Thompson and Bell, length based 5.5
Total length. 1.5
Total mortality coefficient, Z. 4.2
Tr, age at recruitment. 3.1, 4.1
Trash fish. 7.3
Trawl selection 6.1
t-zero, t_o (initial condition parameter). 3.1

Unexploited biomass, Bv 9.1
Unit stock. 1.1

Variance or VAR, s^2 . . . . 2.1
Variance from frequency sample. 2.1
Variance about the regression line. 2.4
Variance of estimates of slope and intercept. 2.4
Virgin stock biomass, Bv. 9.1
Virtual population Analysis (VPA) 5.1
Virtual population Analysis, multispecies 10.2

Wetherall *et al*. method . . . .(3.3.2), 4.5.4
Winterpoint (ELEFAN), tw. 3.5.2
W_∞, asymptotic body weight. 3.3.6

Yield per recruit, Y/R. 8.0, 8.2
Yield per recruit, Beverton and Holt's. 8.2
Yield per recruit, Beverton and Holt's relative 8.4
Yield per recruit, model for a mixed fishery. 10.4.1
Yield per recruit, Thompson and Bell's. 5.3, 10.3

Z, instantaneous rate of total mortality. 4.2

M-43/T0182E/1/9.89/3000
ISBN 92-5-102850-8